Applications of Machine Learning in Aircraft Maintenance

Umur Hüseyin Karaoğlu

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mechanical Engineering

Eastern Mediterranean University
June 2024
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Mechanical Engineering.

Assoc. Prof. Dr. Murat Özdenefe
Chair, Department of Mechanical Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Mechanical Engineering.

Prof. Dr. Qasim Zeeshan
Supervisor

Examining Committee

1. Prof. Dr. Hasan Hacışevki
2. Prof. Dr. Qasim Zeeshan
3. Asst. Prof. Dr. Doğuş Hürdoğanoğlu

ABSTRACT

Aircraft maintenance is a wide-ranging field that involves designing and optimizing various maintenance tasks. This includes finding, detecting, and fixing problems; overhauling, repairing, or modifying aircraft systems; and planning maintenance schedules according to aviation standards. The goal is to predict and prevent failures to keep aircraft reliable. Advances in Big Data Analytics (BDA) and artificial intelligence have changed predictive maintenance. Advances in BDA and artificial intelligence techniques have revolutionized predictive maintenance operations. Predictive maintenance is making big strides in the aviation sector accompanied by a variety of prognostic health management options. AI algorithms have recently been extensively applied to optimize aircraft maintenance systems and operations. Several researchers have proposed, analysed, and investigated the applications of AI, ML and DL based data analytics for predictive maintenance of aircraft systems, subsystems, and components. This thesis provides a comprehensive review of the ML techniques in details. By select three of the ML techniques (ANN, Regression Tree and Linear Regression) from review of ML techniques and comparing that technique to other analysis done by other authors by using same data to be able to maximise efficiency of aircraft maintenance taking account safety, time and cost. It is clearly seen that achieving a 92 percent successful result in one of my models using ANN method which has 80% training data is used with 10-layer size and also, which can be called the best result among the other results. Therefore, I believe it is a promising and improvable method.

Keywords: Artificial Intelligence; Big Data; Deep Learning; Machine Learning; Aircraft Maintenance; Predictive Maintenance; Remaining Useful Life

ÖZ

Uçak bakımı, kapsamlı bakım işlemleri, prosedürlerinin sağlam tasarımını, optimizasyonunu gerektiren geniş kapsamlı ve çok disiplinli bir alandır. Havacılık standartlarına uygun olarak, uçak sistemleri, alt sistemleri, bileşenlerinin arıza tespiti, düzeltilmesi ve revizyonu, onarımı veya modifikasyonunun yanı sıra çeşitli bakım operasyonlarının programlanmasını kapsayan; Arızaları tahmin etmek, önlemek ve böylece uçağın sürekli güvenilirliğini sağlamak. AI tekniklerindeki ilerlemeler, tahmine dayalı bakım operasyonlarında devrim yarattı. Kestirimci bakım, çeşitli prognostik yönetimlerle birlikte havacılık sektöründe büyük ilerlemeler kaydediyor. Yapay zekâ algoritmaları son zamanlarda uçak bakım sistemlerini ve operasyonlarını optimize etmek için kapsamlı bir şekilde uygulanmaktadır. Birçok araştırmacı, uçak sistemleri ve bileşenlerinin tahmine dayalı bakımı için AI, ML ve DL tabanlı veri analitiği uygulamalarını analiz etti ve araştırdı. Bu tez, yapay zekâ tekniklerinin ayrıntılı olarak kapsamlı bir incelemesini sunmaktadır. ML tekniklerinin gözden geçirilmesinden en iyi ML tekniklerinden birini seçerek ve bu tekniği, güvenlik, zaman ve maliyet dikkate alınarak uçak bakımının verimliliğini en üst düzeye çıkarmak için aynı verileri kullanarak diğer yazarlar tarafından yapılan diğer analizlerle karşılaştırılmıştır. Yüzde 80 eğitim verisine sahip, 10 katmanlı boyutun kullanıldığı ve diğer sonuçlar arasında en iyi sonuç diyebileceğimiz YSA yöntemini kullanan modellerimden birinde yüzde 98 başarılı sonuç elde edildiği açıkça görülüyor. Bu nedenle umut verici ve geliştirilebilir bir yöntem olduğunu düşünüyorum.

Anahtar Kelimeler: Yapay Zeka; Big Data; Deep Learning; Machine Learning; Aircraft Maintenance; Predictive Maintenance; Remaining Useful Life

DEDICATION

To My Family and my love,

ACKNOWLEDGEMENT

Human mistake is the most frequent cause of aircraft malfunctions. This can involve a combination of general oversight events, human and design defects in the structure, mistakes made by the flight crew, and improper or misdirected maintenance. Maintaining an airplane involves routinely checking, repairing, and replacing its parts, systems, and structures to make sure it is safe and dependable for flight. The foundation of this study is how maintenance can be maximized while taking cost, time, and safety into account.

I would like to record my gratitude to Prof. Dr. Qasim Zeeshan for his supervision, advice, and guidance from the start of my master's degree. He was my advisor and as the time passed, he became my mentor and always helped me and believed in me. I am very lucky to have a teacher, advisor, mentor, and he had even been a friend for me. His insights, experience, and passion have significantly inspired and enhanced my development as a master's degree student. I owe him more gratitude than he realizes.

I want to thank to my family and my love; they supported me and courage me and never made me feel alone. They always stand by my side and supported my dreams and also, they taught me how to be a good person and how to follow my dreams. I am grateful in every possible way.

I am always going to be proud of being a part of a massive family and I am never going to forget my memories that I had at Eastern Mediterranean University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AE	Acoustic Emission
ANN	Artificial Neural Network
CC	Correlation Coefficient
DRF	Distributed Random Forests
GA	Genetic Algorithms
GBM	Gradient Boosting Method
ILP	Integer Linear Programming
LR	Linear Regression
MAE	Mean Absolute Error
MDP	Markov Decision Process
MLP	Multilayer Perceptron
MRO	Maintenance Repair Overhaul
MTTF	Mean Time to Failure
OAMRP	Operational Aircraft Maintenance Routing Problem
PHM	Prognostics and Health Management
PM	Predictive Maintenance
RF	Random Forest
RL	Reinforcement Learning
RMSE	Root Mean Square Error
RUL	Remaining Useful Lifetime
SVM	Support Vector Machine
SVR	Support Vector Regression
TF-IDF	Term Frequency-Inverse Document Frequency

Chapter 1

INTRODUCTION

1.1 Introduction

Predictive maintenance (PM) involves forecasting maintenance requirements in future using time-based data from in-service facilities such as airplanes. One of the main goals of this method is to accurately forecast when it is time to repair or replace a component. The benefits of prolonged usage of a component diminish if its replacement is done farther ahead; if it's performed very late, unforeseen failures can occur lowering asset availability. As a result, improving the accuracy of estimated component lifetime is a continuous priority. Machine learning (ML) is a subfield of artificial intelligence (AI) in which learning algorithms analyse enormous datasets to discover complex patterns. Despite the fact that the basic architecture and ideas for ML were devised generations earlier, the data volumes and computer power required to make it a reality did not exist until lately. ML applications benefits include maintenance cost reduction, repair stop reduction, machine fault reduction, spare-part life increases and inventory reduction, operator safety enhancement, increased production, repair verification, sensor calibration, fuel tank quantity evaluation, icing detection, and increase in overall profit, and many more (Çinar et al., 2020; Cline et al., 2017; Patrick KY, 2020).

For far more than a century, aerospace has relied on "preventative" maintenance, which includes simplistic analytical expression that avoid attempting to predict failures in favour of erring on the side of caution and replacing parts in a systematic way, using many equipment sensor readings as a guideline marker of system health. Predictive machine learning algorithms, on either side, can account for complicated underlying relationships by using knowledge embedded in previously untapped datasets. As a consequence, such systems can assist in identifying failure patterns in components that would be difficult to discover otherwise.

To begin with, having a vast volume of data is insufficient. Defining all of the parameters that potentially influence equipment behaviour is required when developing a machine learning model. This can be achieved through cooperation between data scientists and experienced field engineers, as well as a desire to comprehend each other's environment and perspective. Second, data must be collected and "cleaned" to focus on the variables that affect equipment performance, such as operating circumstances, temperature exposure, and previous repairs. However, data does not need to be flawless; it was discovered that freely available data can be used to begin training a machine learning model and improve prediction accuracy. Finally, the appropriate machine learning model must be chosen and developed for the task. Systems get more effective when they become harder to comprehend. So, a simple "linear" machine learning model will produce outputs that are simpler to use (though not as accurate), whereas complicated "neural networks" would yield higher accuracy, however the algorithms itself are much more tough to comprehend.

1.2 Motivation

Maintaining an aircraft in airworthy condition that is, ensuring it is safe to fly and satisfies the performance criteria specified by aviation authorities is the primary goal of aircraft maintenance. This entails routinely inspecting a variety of parts, including the engines, avionics, airframe, landing gear, hydraulic and electrical systems, and other parts, to make sure everything is in working order and is free from damage or faults. There are two primary types of aircraft maintenance: planned maintenance and unplanned maintenance. Scheduled maintenance is performed in accordance with a planned maintenance program that takes into account the aircraft's operating hours, cycles, or calendar time. It consists of routine inspections, preventive maintenance, and overhaul processes. This kind of maintenance minimizes the chance of malfunctions or breakdowns and is planned and scheduled ahead of time. Conversely, unscheduled maintenance is carried out as necessary and is typically brought on by unforeseen issues or flaws that develop while the machine is in use. Unscheduled maintenance includes changing worn-out or damaged parts and fixing issues with the engine, hydraulic, or electrical systems. In order to stop additional damage or safety hazards, this kind of maintenance is frequently more essential and needs to be done right away. The foundation of this study is how maintenance can be maximized while taking cost, time, and safety into account.

1.3 Scope and Objectives

This study aims to contribute recent researches about aircraft maintenance giving the idea of recent researches and the algorithms used. The main objective is the algorithm that is used in this research is the algorithm that never used before and compared to other results. Also, that algorithm is modelled differently for best result. For example, percentage of training data is used, percentage of the validation data, percentage of test data or layer size. Also, the data that is selected plays an important role on this research by selecting an aircraft data that is available gives this research better result. As internal implementation details are under- documented in the literature, the present study provides an insight to the aircraft maintenance by involving machine learning algorithm.

1.4 Thesis Outline

Chapter 2 Discusses history and past researches about aircraft maintenance and machine learning techniques.

Chapter 3 A literature review of aircraft maintenance, machine learning implementation, parameters, objectives, constraints, and simulation programs is presented.

Chapter 4 The methodology of how the data is used and how the analysis is done is briefly explained.

Chapter 5 Results of analysis and further comparison with results from other sources is presented for validation.

Chapter 6 The findings are concluded and a future outlook is presented.

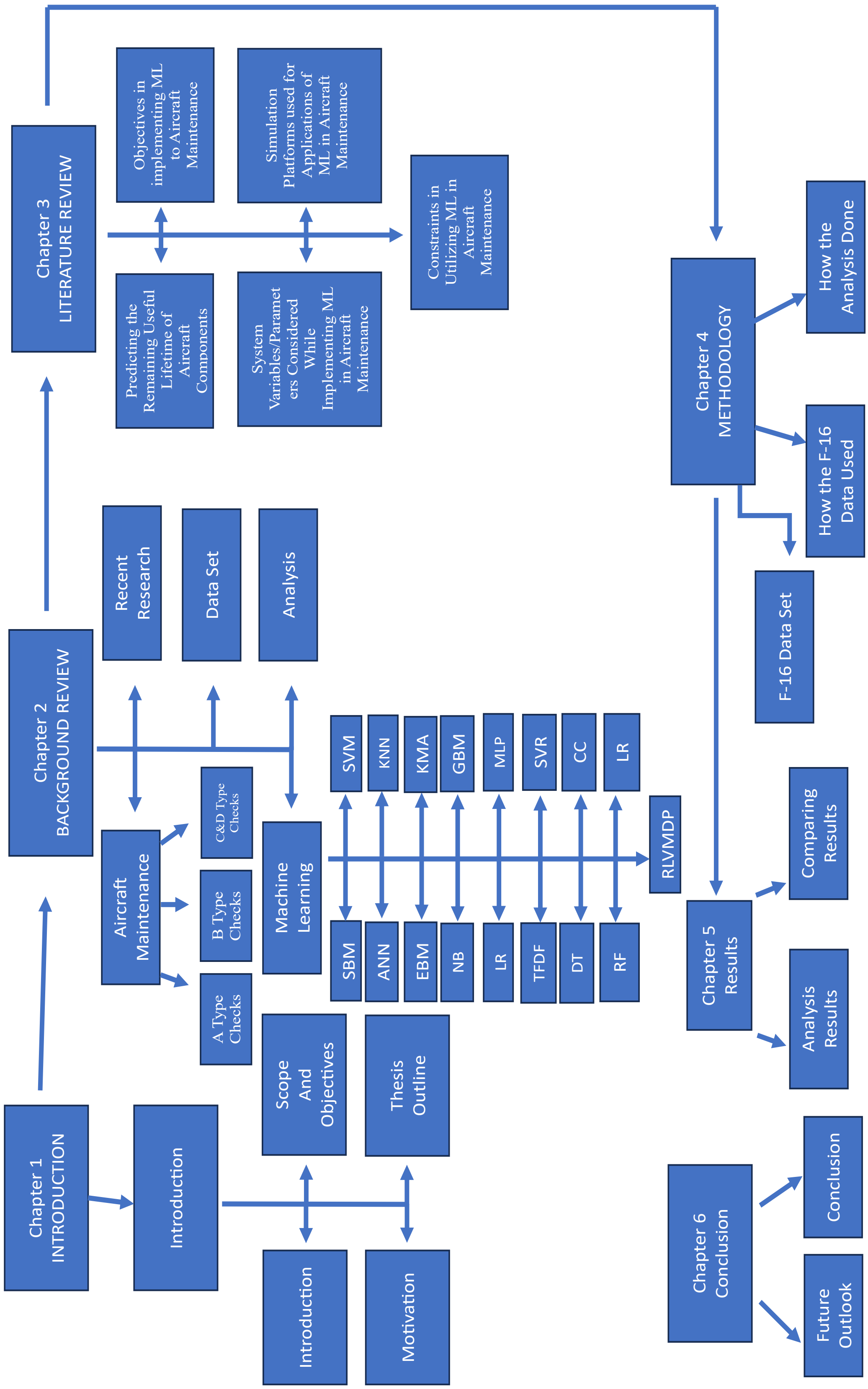


Figure 1: Flow Chart

Chapter 2

BACKGROUND INFORMATION

The thesis is about how to use machine learning algorithms precisely by comparing other results to be able to decrease cost and time by increasing safety of aircraft maintenance. F-16 Data set is chosen as well as Artificial Neural Network will be handled.

2.1 Aircraft Maintenance

The execution of duties necessary to guarantee an aircraft's or an aircraft part's continued airworthiness, such as overhaul, inspection, replacement, defect rectification, and the incorporation of modifications, compliance with airworthiness directives, and repair, is known as aircraft maintenance. Parameters like safety, time and cost is very important during aircraft maintenance. Apart from unexpected repairs, aircraft maintenance is performed in a sequence of more thorough checks. How frequently these checks are carried out relates to the number of flying hours, take-off and landing rotations, and can be performed at any location with the necessary equipment. Since each aircraft type has its own inventory requirements, integrating facilities for various fleets saves just a little amount of money. Several companies have developed maintenance practices that require periodic checks at least every four days to comply with Federal Aviation Administration (FAA) requirements in United States. The FAA requires each aircraft to go through four main types of inspections. The range, timing, and regularity of these events vary (Clarke et al., 1997).

2.1.1 A Type Checks

The FAA requires the first main check (A Type) per 65 hours of flight, or approximately once in a week. During Type A inspections, critical systems like engines, landing gears, and control surfaces at the cockpit are inspected (Hobbs, 2021).

2.1.2 B Type Checks

Every 300-600 flying hours, B Type which includes a comprehensive visual inspection together with greasing all moving parts such as ailerons and horizontal stabilizers is carried out (Hobbs, 2021).

2.1.3 C&D Type Checks

The Type C and D primary inspections are performed once in a four-year period and necessitate the aircraft to be out of use for at least a month. Type C and D inspections don't have to be considered in maintenance scheduling since they are spread at relatively long periods and also because of the dynamic nature of the industry. The airlines' main focus is fulfilling the Type A and B inspection and maintenance criteria in their chosen 4-day inspection and maintenance program. Inspections and repairs are done at night until they are in unique conditions (Hobbs, 2021).

2.2 Recent Research

The problem of aircraft maintenance management is highly complicated, encompassing a variety of factors that can be divided into three major sub-problems: project scheduling and resource allocation for scheduled and unscheduled duties, as well as personnel capacity planning, spare parts forecasts, and stock control (Dinis & Barbosa-Póvoa, 2015). In this sense, capacity planning refers to the practice by which Maintenance Repair Overhaul (MRO) businesses allocate their workforce at a given point in time to meet future demand. Dijkstra et al. (1991) reported their work at KLM, the Royal Dutch Airlines, on designing a Decision Support System (DSS) to solve capacity planning problem. Because the topic is NP-hard, the optimization model is presented an integer programming and is dealt with using an approximation approach based on Lagrange relaxation. Data related to labour and manpower is stored in a database module, a methodology assesses potential outcomes in terms of scenarios, and DSS is contained in a graphical user interface (GUI). The methodology which makes use of optimization models contains routines for estimating labour, optimizing workforce size, organizing and evaluating computed labour's quality (Dijkstra et al., 1994). Evaluations routine aims at maximizing the total number of works to be completed with the designated size and shape of labour, whereas the optimization model's routines aim to achieve the least number of maintenance engineers necessary to finish all jobs. S. Yan et al. (2004) published an airplane maintenance capacity planning model that incorporates different maintenance certificates and flexible management tactics. In increasing the various degrees of flexibility in the maintenance schedules, S. Yan et al. (2004) suggested three flexible solutions. (1) flexible shifts, that gives room for the maintenance management to schedule the best values of shift patterns and their continuous phase; (2) flexible teams (also known as squadrons), which allow companies to modify the number of team members in terms of supply; (3) flexible working hours, that enables the firm to create separate number of hours. The purpose of the challenge which is modelled as a mixed integer algorithm, is to decrease overall maintenance manpower (man-hour) while still fulfilling demand.

Regattieri et al. (2005) highlighted two methods to select spare parts of aircrafts: (1) selection based on operating expertise of an enterprise; and (2) the use of prediction methodologies. Comparison was made on the performance of twenty prediction methodologies using some dataset from Alitalia airline in forecasting the need for aircraft spare parts in case of irregular demands of spare parts and it was observed that Croston method, exponentially weighted moving average, and weighted moving averages models outperformed the rest when evaluated with mean absolute deviation (MAD). As evidenced by some researchers (Ghobbar & Friend, 2003; Regattieri et al., 2005; Romeijnders et al., 2012) in predicting aircraft spare parts, the latter mostly made use of statistical estimate techniques such as the Croston's approach, single exponential smoothing (SES), and the simple moving average (SMA). Neural Networks (NNs) was used in making prediction for irregular and lumpy demand patterns (Gutierrez et al., 2008; Kourentzes, 2013), although not explicitly in aviation. Optimization methodologies have not previously been used in combination with spare parts predictions to the authors' knowledge. They are, nevertheless, frequently used to solve stock management issues. In order to address production/inventory dilemma with non-stationary supply unpredictability and the accessibility of reasonably close supply information (called advance supply information), (Atasoy et al., 2012) offered a dynamic programming technique. W. Wang et al. (2012) combined a dynamic programming and an enumeration coding approach to enhance the spare parts inventory using regular preventive maintenance period. The subject of managing logistics of spare parts and its operations for maintenance service suppliers was addressed by Kazemi Zanjani & Nourelfath et al. (2014). The best combination of maintenance tasks that can be carried out and the number of spare part orders that will reduce the costs of inventory, procurement, and late delivery while taking into consideration the lead time for supplying spare parts was first determined by Kazemi Zanjani & Nourelfath et al. (2014) using a theoretical mathematical programming model. Also, the non-stationary Variable model which is later reconstructed as a multi-stage random programming with access was used to represent the request uncertainty for spare parts. Gu et al. (2015) presented two non-linear modelling techniques for deciding the best procurement duration and quantity while keeping costs to a minimum.

Huang et al. (2010) used a GA to allocate human resources to an problem involving airplane maintenance. They started with a specified sequence of maintenance chores and then figured out the best distribution of resources to keep the maintenance tasks as short as possible. Safaei et al. (2011) focused on issues regarding scheduling maintenance tasks for a fleet of military aircraft. These problems which was constructed as a mixed-integer programming model (MIP) was solved using a branch-and-bound approach. Maximizing fleet availability was the key objective of the model, but the availability of skilled labour was the main.

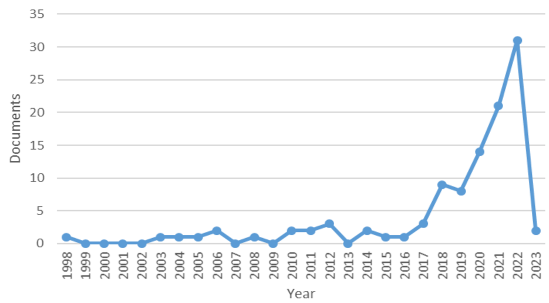
Nuhu et al. (2023) investigated the applicability and effectiveness of ML prediction models for fault diagnosis in smart manufacturing. In predicting the Remaining Useful Lifetime (RUL) of aircraft parts, Azevedo et al. (2019b) developed a web-based application. This application models a Prognostics and Health Management (PHM) system with an aim of forecasting the RUL of specific aircraft parts by implementing various machine learning algorithms such as these three proposed methods: extrapolation-based method, similarity-based method and neural network-based method. MATLAB 2018a was used to implement the three techniques. Nguyen & Medjaher et al. (2019) designed a dynamic PdM structure based on LSTM, a deep learning architecture, to look into the impact imperfect prognostics have on maintenance decisions in Prognostics and Health Management (PHM) system. Sarkon et al. (2022) reviewed ML applications in additive manufacturing; from design to manufacturing and property control. T. Wang et al. (2008) suggested a prognostics approach based on similarity for estimating the RUL of designed systems, of which a unique data-driven RUL estimation methodology that begins with damage assessment and then employs a similarity-based matching method to compute RUL was presented. The method is made up of two procedures: RUL estimate and performance evaluation.

Capodieci et al. (2020) employed various machine learning algorithms to evaluate spare parts request performance for aircraft engine by developing iterative Artificial Intelligence-based algorithms to outline the plan for engine removal and maintenance, enhance maintenance costs and engine availability at the customer, and also obtain an acquisition plan of integrated parts with intervention planning and maintenance scheme execution. Machine Learning was used on a workshop dataset to optimize the amount, cost, and lead-time of warehouse spare parts in order to achieve this goal. This dataset comprises of information such as repair claims, engine operating hours, forensic evidence, and general information regarding processed spare parts for a given engine type, spanning various years and fleets. The Confusion Matrix was then used to assess the overall results and make comparison using three machine learning algorithms - Naive Bayes, Logistic Regression, and Random Forest classifiers. Naive Bayes and Logistic Regression estimates perform best globally, with an accuracy rate of roughly 80%, indicating that the models are generally accurate.

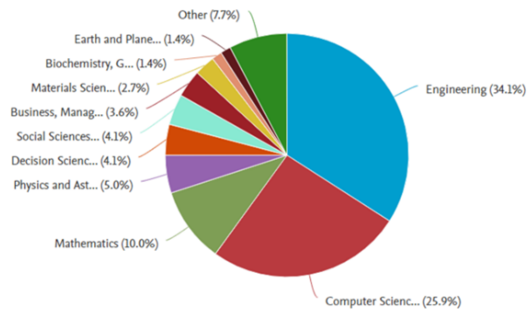
W. Yan & Zhou (2017) used term frequency-inverse document frequency (TF-IDF) and random forest to develop a predictive model that forecasts high-priority defects in advance in order to carry out preventive maintenance. This model was built on historical data from airplane maintenance systems. W. Yan & Zhou (2017) formulated the prediction of faults with varying importance as a binary classification problem. Counts of different defects that have occurred in previous flights was used as raw data. TF-IDF algorithm was then implemented in extracting features from the raw data. The RF method was used to simulate the classification of defects with varied priorities. The training dataset's performance measurements are based on the Receiver Operating Characteristics (ROC) curve. Mathew et al. (2017) performed a comparative analysis on existing machine learning techniques for forecasting the RUL of an aircraft's turbofan engine. These techniques are K-Means algorithm, Linear Regression, Random Forest, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbours, Gradient Boosting Method (GBM), and AdaBoost. Research carried out by El Afia & Sarhani (2017) which is about model selection of aircraft maintenance predictive models using particle swarm optimization was tested in the calculation of an aircraft's RUL, which has an impact on maintenance planning. Autoregressive integrated moving average (ARIMA) models are the most common predictive models, and they're extremely useful for forecasting. Machine learning techniques such as ANN and SVM are used to create predictive models. Azevedo et al. (2019a) developed an online simulation tool to allow users simulate a Prognostics and Health Management (PHM) system, build and post machine learning exploratory models in order to predict the RUL of aircraft parts impacted by a system fault with Neural Network based method, Similarity based method and an Extrapolation based method. Partha Adhikari et al. (2018) developed a Data Driven Diagnostics & Prognostics Structure to perform aircraft predictive maintenance based on machine learning. Deng & Santos et al. (2022) proposed a Lookahead approximate dynamic programming framework (incorporating a hybrid lookahead scheduling policy) which uses deterministic prediction to make the best decision for heavy maintenance of aircraft and stochastic prediction to decide light airplane maintenance. Due to unpredictability of aircraft's daily operation and maintenance check elapsed time Deng & Santos (2022)'s work aimed at reducing the wasted time period between maintenance checks. Celikmih et al. (2020) utilized machine learning algorithms to develop a feature selection and data elimination model in order to forecast aircraft equipment failures. Using aircraft maintenance and failure data collected over a period of time, this model was able to separate important parameters from ineffective ones in the data. K-means algorithm was then implemented in the second step to remove inconsistency in the data. Multilayer Perceptron (MLP) which is an Artificial Neural Network (ANN), and Support Vector Regression (SVR), and Linear Regression (LR) which are machine learning techniques were used on the equipment maintenance dataset to evaluate the performance of the model. Also, the model was evaluated using performance metrics such as the Correlation Coefficient (CC), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Results showed that the hybrid data preparation model is effective at forecasting the equipment failure rate. Ai et al. (2021) showed how machine learning algorithms can be implemented to detect impacts on aircraft composite structures which describes a promising technique for automatically detecting and localizing a debris or hail impact during flight. Acoustic emission (AE) was used as an impact surveillance method to attain this purpose. The flowchart of this study is shown in Figure 3. An overview of the implementation of various deep learning techniques used for aircraft maintenance, repair, and overhaul (MRO) was conducted by (Rengasamy et al., 2018) . Four architectures found to have been used for MRO were Deep Autoencoders, Convolutional Neural Networks (CNN), Deep Belief Networks and Long Short-Term Memory. S. Paul (2010) proposed an MRO system them applies ANN to eliminate deficiencies in existing aircraft MRO systems.

Novel strategies based on machine learning models to score rotorcraft maintenance logbook event data for reporting purposes was developed by (Seale et al., 2019). Manual inspection of huge generated maintenance has led to about 10% of the data receiving scores from human analyst. Machine learning algorithms are introduced to solve this but despite the fact that the computing efficiency of standard classification algorithms has significantly improved, using these machine learning methods to solve issues with huge numbers of unique class labels remains difficult. Using strategic label set segmentation and hierarchical ensemble models, (Seale et al., 2019) employed distributed random forest to handle this challenge. Antolick (2015) designed two decision making models for engine output gearbox and turbo-shaft engine using data obtained from health and usage monitoring systems (HUMS). Support vector machine (SVM) was used to create decision boundaries on this dataset and receiver operating characteristic (ROC) curve was used to analyse the simulated work and observe the performance of the employed machine learning method. Dangut et al. (2021) proposed a hybrid model for aircraft parts rare failure prognostics. This hybrid model which combines natural language processing techniques with ensemble learning is used for forecasting rare airplane parts failure. The model is validated with a genuine aircraft central maintenance system log-based dataset. The performance of various machine learning algorithms implemented in the reliability analysis of aircraft engine was investigated by Singh et al. (2020), and a superior approach for predicting RUL was proposed. Different classification and regression algorithms were used. Examples of classification algorithms were Ensemble method, Random Forest classifier, CatBoost classifier, Light GBM classifier and XGBoost classifier and examples of Regression Algorithms were Random Forests Regression, Ridge regression, Lasso regression, XGBoost regression, Elastic-net regression, Support vector Machine as Regression, K-Nearest Neighbour Regression, Decision Tree Regression, Adaboost Regression and Linear Regression. Andrade et al. (2021) improved long-term maintenance check scheduling for aircraft fleets using Reinforcement Learning (RL). Scheduling hangar checks for a given time horizon takes into account the fleet state, repair capacity, and other maintenance limitations. The inspections are planned at regular intervals, with the intention of scheduling them as close as possible to their due date.

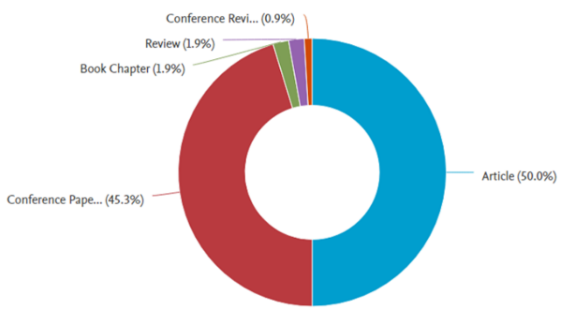
Airline companies with a drive for profit making have begun to concentrate on getting the best routes that are viable to maintain. Operational aircraft maintenance routing problem (OAMRP) was investigated by Ruan et al. (2021) with the goal of firstly offering an Integer Linear Programming (ILP) framework for the OAMRP that simultaneously analyse these three key maintenance constraints: maximum take-offs between two successive maintenance checks, maximum flying-hour, and capacity of work-force, and secondly designing a novel reinforcement learning-based algorithm that will provide a fast and efficiently solution to the problem. Testing this approach on a dataset received from a prominent Middle Eastern airline showed that it produces outstanding solutions for both datasets of medium and large-scale flight schedule. The Markov Decision Process (MDP) is a problem that can be handled utilizing an existing reinforcement learning-based method to develop best routes which are also maintainable. An earlier work published by (Doğru et al., 2020) showed the implementation of MASK R-CNN, a modern Convolutional Neural Network architecture, on autonomous drones to assist aircraft maintenance engineers to detect faults in airplanes. The reason for the choice of MASK-RCNN as seen in Bouarfa et al. (2020) was because it detects several objects in an image while producing a segmentation mask for each occurrence. Doğru et al. (2020) built on their earlier work by experimenting various ways to evaluate prediction and make it more accurate. These experimented ways included (1) focusing only on wing images to increase data homogeneity; (2) adding images without dents to balance the original dataset; (3) investigating the potential of some augmentation techniques which are rotating, flipping, and blurring, in improving model performance; and (4) using a pre-classifier in conjunction with MASK R-CNN. Figure 1 shows the statistics of articles published over the years between 1998 and 2023. These statistics is obtained from Scopus database using (TITLE-ABS-KEY ("aircraft maintenance") AND ("machine learning")) as the search keyword. The number of published articles using major keywords of this review is shown in Figure 2.



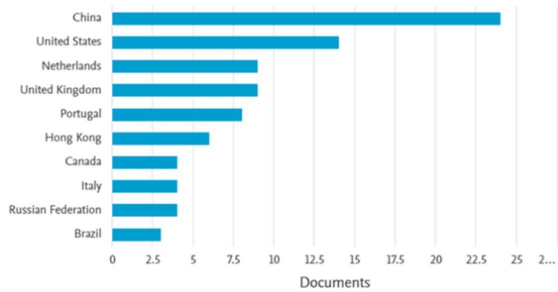
(a)



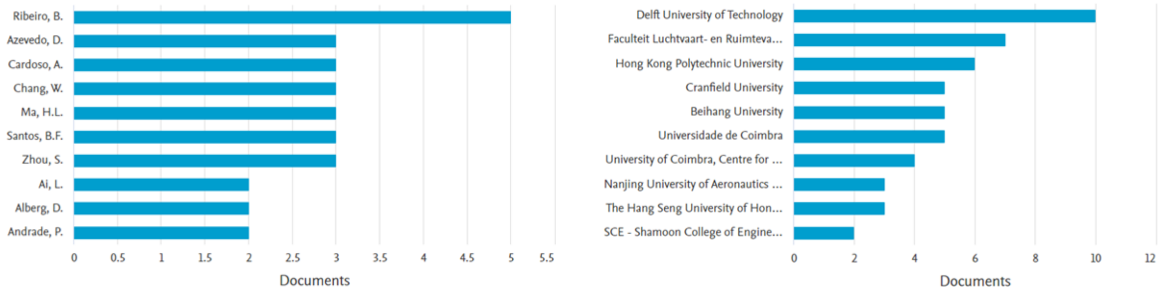
(b)



(c)

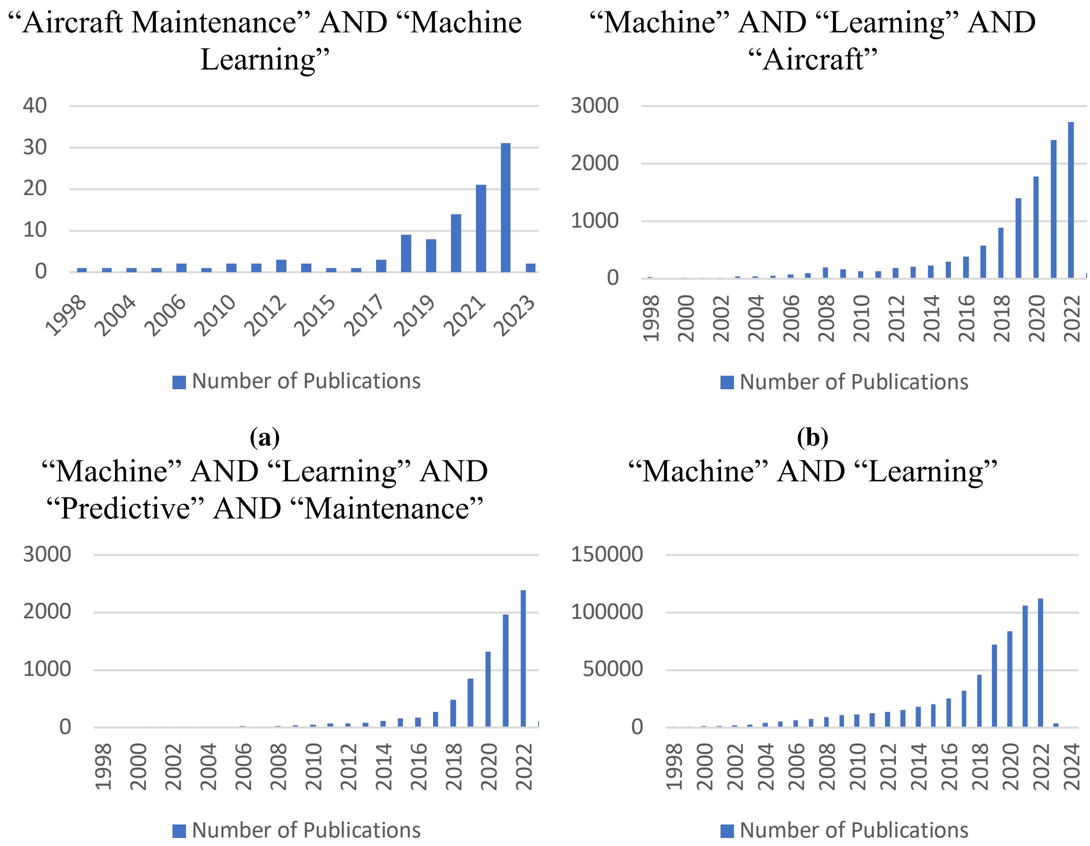


(d)



(e) (f)

Figure 2: Statistics from Scopus Database



(a) (b) (c) (d)

Figure 3: Statistics from Scopus Database Research

2.3 Machine Learning

In artificial intelligence, machine learning (ML) is the study and creation of statistical algorithms that can learn from data, generalize to new data, and carry out tasks without explicit instructions. Artificial neural networks have recently demonstrated performance gains over several earlier methods. A subset of machine learning models called neural networks also referred to as artificial neural networks, neural nets, or simply ANN or NN are modelled after the neuronal architecture of biological neural networks seen in animal brains.

2.3.1 Classification of Machine Learning Methods

2.3.1.1 Similarity Based Method

SBMs are a generalization of minimal distance (MD) approaches, which are used in a variety of machine learning and pattern recognition algorithms (Azevedo et al., 2019b; T. Wang et al., 2008). A nearest neighbor approach is used in similarity, a machine learning technique, to compare two or more objects using algorithmic distance functions (Azevedo et al., 2019a; I. Dave, 2018).

2.3.1.2 Neural Network

Different neural network types such as artificial neural network (ANN) and convolutional neural network (CNN) operate according to various principles and establish their own rules. A neural network (NN) is artificial neuronal circuit (Azevedo et al., 2019b). The connection of neurons is modelled by ANN as weights between nodes. Excitatory connections have a positive weight, while inhibitory connections have a negative weight (El Afia & Sarhani, 2017). Each input is given a weight and then summed. This is termed linear combination. Lastly, an activation function regulates the output's amplitude. Output range can be in the interval [0, 1], though it may also be in the interval [-1, 1] (Azevedo et al., 2019a). Figure 4 shows the structure of a neural network. ANN is used for applications that require training a dataset such as predictive modelling, and adaptive control. Networks can learn from themselves based on experience therefore, drawing conclusions from a large and seemingly unconnected set of data (Celikmih et al., 2020). A CNN is a type of NN used in recognising and processing images (Doğru et al., 2020; Rengasamy et al., 2018). ANN is a massively parallel computing system that consist of large number of basic processors linked together by a large number of interconnections. ANNs learn the basic principles from a series of supplied symbolic scenarios or try to understand patterns in a dataset to make predictions, rather than following a set of laws established by human experts. Furthermore, the source of these ANNs' analytical activity is the relationships between the network processing units.

2.3.1.3 Extrapolation Based Method

Extrapolation is a method for estimating the value of a variable beyond the initial observation range based on its relationship with another variable (Azevedo et al., 2019b). Extrapolation is similar to interpolation in that it provides estimates between known observations, but it is riskier and more likely to produce meaningless results. Extrapolation can also refer to the expansion of a method based on the assumption that similar approaches are applicable (Azevedo et al., 2019a). Extrapolation is the process of projecting, extending, or expanding known knowledge into an unknown or previously unexperienced area in order to gain insight into the unknown. Extrapolation is a technique for estimating a value that lies outside of a training data region (McCartney et al., 2019).

2.3.1.4 Naive Bayes

It's a classification method built on Bayes' Theorem (Capodieci et al., 2020). Naive Bayes classifier postulates that a feature existing in a class has no bearing on the presence of any other features. The Naive Bayes model is easy to develop and performs better with large data sets. Also, it can be used to detect and reduce noise in data (Jalawkhan & Mustafa, 2021).

2.3.1.5 Logistic Regression

Logit model which is a statistical analysis tool is used for predictive analytics and modelling. In this analytics approach, dependent variables are either finite or in categories i.e. either a binary regression (A or B) or a multiple/multinomial regression (A, B, C, or D). Logistic regression helps to understand the connection between a dependent variable and one or more independent variables by estimating probabilities (Capodieci et al., 2020). Analysis of multiple variables with this tool helps to reduce confounding factor in variables (Sperandei, 2014).

2.3.1.6 Decision Tree (DT)

DT is a type of probability tree that aids making a choice about a particular process. For instance, one could choose whether to produce item A or item B (Mathew et al., 2017). In dealing with complicated decisions with lots of variables and are usually ambiguous, DT is a great approach. Since trees can quickly grow complex, using a software is usually preferred even though they can be drawn by hand (Luna et al., 2019). A Decision Tree is a network structure made up mostly of nodes and branches, with root nodes and intermediate nodes constituting the nodes. The leaf nodes indicate a class label, whereas the intermediate nodes contain a characteristic. DT classifiers have grown in prominence in a variety of fields, including character recognition, medical diagnosis, and voice recognition (Ahmad et al., 2019; Akamine & Ajmera, 2012; Grafmüller et al., 2010).

2.3.1.7 Random Forest (RF)

Random forest, developed by Breiman et al. (2001) employs an ensemble learning method for solving complex problems by merging multiple classifiers (Capodieci et al., 2020; Seale et al., 2019). RF algorithm consist of many decision trees which can be trained using bagging or bootstrap aggregation. Bagging is a method used to increase their accuracy (Altman & Krzywinski, 2017). RF algorithm determines the outcome from predictions of DT by averaging their results. As the number of trees grows result's accuracy improves. The output category is decided collectively by these individual trees in an ensemble learning process, which is made up of numerous DT classifiers. Generalization error converges as the number of trees grow. The RF has a number of other advantages. For example, without selecting a feature, it may handle high-dimensional data. Trees are independent of one another throughout the training process, and implementation is very easy, yet, training speed is normally rapid, and generalization functionality is enough.

2.3.1.8 Term Frequency-Inverse Document Frequency

How frequently a phrase appears in a document is indicated by term frequency (TF). Terms and sentences or words are synonymous in natural language. But any text token can be represented by a word. A phrase is probably going to appear more frequently in longer documents than in shorter ones as documents come in all different lengths (W. Yan & Zhou, 2017). As a result, a term will appear to be more essential in a longer text than in a shorter one. Term frequency is regularly divided by the total number of terms in the document in order to normalize the effect. TF checks the frequency of a term *t* in document *d*, while DF counts the occurrence of term *t* in document set *N*. DF stands for the number of documents in which the word appears (G. Kavita, n.d.).

2.3.1.9 Linear Regression

A linear model tells a linear relationship between input and output variable x and y (Mathew et al., 2017). y is determined by a linear combination of the input variables x. For a single input variable x, this procedure is known as linear regression while several input variables describe multiple linear regression in statistics literature (Clarke et al., 1997). Linear regression is a multivariate linear combination of regression coefficients. The generalized least square approach is used to determine the coefficients. Linear regression has its application in time series regression techniques that have been employed in predictive maintenance (Zenisek et al., 2019).

2.3.1.10 Support Vector Machine (SVM)

SVM is a linear model that can be used to solve classification and regression issues Csurcsia et al. (2020) and also linear and nonlinear problems El Afia & Sarhani (2017). SVM divides data into classes by drawing a line or hyperplane (Blanco et al., 2018). A statistical learning idea with an adaptive computational learning approach is defined as SVM. SVM, a supervised machine learning approach, can achieve pattern classification, recognition, and regression analysis. SVMs have been frequently used in the PdM of industrial equipment to determine a certain condition based on the recorded signal (Widodo & Yang, 2007).

2.3.1.11 K-Nearest Neighbours

The nearest neighbour (NN) rule is a technique for classification. It categorizes a sample based on its closest neighbour's category (Mathew et al., 2017). The probability of error in NN is less than twice the optimum error when many samples are used thus having a lower probability of error than any other decision rule. To categorize a test pattern, nearest neighbour-based classifiers employ part or all of the patterns in the training set. The fundamental goal of these classifiers is to find correlations between each pattern in the training set and the test pattern (Murty & Devi, 2011).

2.3.1.12 The K-Means Algorithm

The K-means clustering technique, also known as Flat clustering algorithm, calculates centroids and iterates to get the best centroid. It speculates the number of clusters that exist (Mathew et al., 2017). 'K' in K-means is the number of clusters identified by the algorithm from the data. The data points are assigned to clusters in such a way that the sum of their squared distances from the centroid is as small as possible (Celikmih et al., 2020). Various distance metrics such as Euclidean, Manhattan, and Minkowski can be employed to observe the behaviour of clustering results (Ghazal et al., 2021).

2.3.1.13 Gradient Boosting Method (GBM)

Gradient boosting is used in classification and regression tasks (Mathew et al., 2017). It is a group of sophisticated machine-learning algorithms (XGBoost, LightGBM and CatBoost) that have had a lot of success in a variety of applications (Bentéjac et al., 2021). Additionally, it is an ensemble-based model that continuously improves the prediction accuracy of new models (Luna et al., 2019).

2.3.1.14 Multilayer Perceptron (MLP)

A Multilayer Perceptron is a fully connected multi-layer ANN. It contains three levels in which one is hidden (Celikmih et al., 2020). MLP uses backpropagation in updating the weight parameters by finding the gradient of the loss function. Training of this ANN starts with random initialization of weight and bias parameters. Training the network aims at developing a model that can generalize well on input data and produce a high accuracy on predicted results (Sano & Berton, 2022).

2.3.1.15 Support Vector Regression (SVR)

Support vectors are points that are outside the tube in SVR. The smaller the value, the more the points outside the tube, and hence the more the support vectors. The name SVM, or Support Vector Machine, is well-known among those who work in Machine Learning or Data Science. SVR, on the other hand, is not the same as SVM (Celikmih et al., 2020). As the name implies, SVR is a regression algorithm, which means we can use it instead of SVM for working with continuous value (Ghobbar & Friend, 2003).

2.3.1.16 Correlation Coefficient (CC)

Correlation coefficient measures the strength of the linear connection between two variables. Correlation coefficient is represented by *r*. The formula calculates the distance between each data point and the variable mean for two variables and uses the result to assess how well the relationships between the variables can be fit to a hypothetical line through the data (Celikmih et al., 2020). Correlation does not deal with connections in bivariate data, rather it only considers the two variables in question. Outliers in the data will not be detected, so it’s important to inspect plotted data to ensure data is within the required range in a fairly uniform way (Asuero et al., 2006).

2.3.1.17 Reinforcement Learning via Markov Decision Process

Reinforcement learning employs Markov Decision Process (MDP), a mathematical framework to describe an environment (Ruan et al., 2021). RL is a branch of machine learning that covers different strategies for learning an optimal policy (i.e., a mapping from states to actions) for an agent in a given environment. The MDP is a method for defining an unknown environment in which the agent takes a couple of actions to maximize its reward (Razzaghi et al., 2022).

Table 1: Classification of Machine Learning Method

Model	References	Remarks
Similarity Based Method	[(Azevedo et al., 2019b), (T. Wang et al., 2008), (Azevedo et al., 2019a), (Andrade et al., 2021), (I. Dave, 2018)]	• Due to the transitive quality of pairwise similarities, they are robust. • They can exploit the pairwise relationships between unlabelled objects to their advantage. • They can interpret data by grouping or labelling raw input. • They can be thought of as a categorization of the grouping layer that sits on top of the stored and managed data.
Neural Network	[(Azevedo et al., 2019b), (El Afia & Sarhani, 2017), (Azevedo et al., 2019a), (Celikmih et al., 2020), (Rengasamy et al., 2018), (Andrade et al., 2021), (Brosnan Yuen, 2020),(Charlotte Daniels, 2020),(Elettronici Course & Tania Cerquitelli, 2017)]	• Neural networks are a component of broader machine learning as services and applications, which includes algorithms for classification, regression, and reinforcement learning. • Enables organizations to make predictions based on the data they have.
Extrapolation Based Method	[(Azevedo et al., 2019b), (Azevedo et al., 2019a), (McCartney et al., 2019)]	• Data requirements are minimal for this forecasting approach. • It's not necessary to collect a lot of data to forecast future data points. • It is simple and straightforward to apply.
Naïve Bayes	[(Capodieci et al., 2020), (Jalawkhan & Mustafa, 2021)]	• It doesn't necessitate as much data for training. • It is capable of dealing with both continuous and discrete data. • It provides fast response when making predictions in real-time.
Logistic Regression	[(Capodieci et al., 2020), (Sperandei, 2014)]	• Training and analysing with Logistic regression is more straightforward. • Logistic Regression cannot be used if the number of data points is less than the list of features; therefore, it may result in overfitting. • It classifies unfamiliar records fairly quickly. • It reduces the effects of confounding factors in analysed variables.
Decision Tree	[(Mathew et al., 2017)]	• A benefit of decision trees is that statistical expertise is not needed in analysing its results. • It's simple to put together. • Data cleansing isn't as necessary. • Multiple problems can be analysed and solved by decision trees. • It performs both regression and classification.
Random Forest	[(Capodieci et al., 2020), (Mathew et al., 2017), (Seale et al., 2019), (Singh et al., 2020), (Dos Santos et al., 2017)]	• It generates accurate predictions that are easy to comprehend. • Can manipulate huge datasets effectively. • It reduces overfitting in decision trees prediction accuracy. • It's capable of working with category and continuous data.

Model	References	Remarks
Term Frequency-Inverse Document Frequency	[(W. Yan & Zhou, 2017), (G. Kavita, n.d.)]	• Simple to calculate. • It knows how to extract the most descriptive terms from a document using some simple metrics. • It can quickly calculate the similarity of two documents.
Linear Regression	[(Mathew et al., 2017), (Singh et al., 2020), (Jason Browniee, 2023)]	• This algorithm is best used when the independent and dependent variables have a linear relationship. • Linear regression is simple to use, and the output coefficients are easier to understand. • Linear regression models can be trained efficiently on systems with little processing resources.
Support Vector Machine (SVM)	[(Mathew et al., 2017), (El Afia & Sarhani, 2017), (Cortes & Vapnik, 1995)]	• Performs efficiently on data that contains a clear margin of distinction among classes. • When the number of dimensions is more than the number of samples, SVM is effective. • SVM takes little memory space when training a model.
K-Nearest Neighbours	[(Mathew et al., 2017), (Nurhadiyanto et al., 2019)]	• Computational time is short. • Results are simple method to evaluate. • It's versatile, and it can be used for regression and classification. • Provides high precision results.
The K Means algorithm	[(Mathew et al., 2017), (Ghazal et al., 2021)]	• Less complicated to implement. • It is capable of manipulating big data sets. • Convergence is guaranteed. • Generalizes to other shapes and sizes of clusters, such as elliptical clusters. • It is reliant on starting values.
Gradient Boosting Method (GBM)	[(Mathew et al., 2017), (Bentéjac et al., 2021), (Essam Al Daoud, 2019)]	• Frequently gives exceptional forecasting accuracy. • Many different loss functions can be optimized, and there are various hyper parameter adjustment possibilities that make the function fit quite versatile.
Multilayer Perceptron (MLP)	[(Celikmih et al., 2020)]	• Works well with large amount of data. • Works with linear and non-linear models. • It has real-time data learning capability.
Support Vector Regression (SVR)	[(Celikmih et al., 2020), (Widodo & Yang, 2007), (Cortes & Vapnik, 1995), (Awad & Khanna, 2015)]	• It can tolerate outliers. • Update the decision model is easy. • It has a high prediction accuracy and generalizes well. • It is simple to implement.
Correlation Coefficient (CC)	[(Celikmih et al., 2020), (Asuero et al., 2006)]	• It can be used to demonstrate the strength of a relationship between two variables. • Obtain quantitative data that can be analysed quickly.
Reinforcement Learning via Markov Decision Process	[(Ruan et al., 2021), (Razzaghi et al., 2022)]	• It enhances performance. • Change can be sustained for a long time.

2.4 Data Set

The data set includes information on the installation of fake payloads on the wings to mimic the characteristics of actual equipment. The aircraft structure was instrumented with 145 acceleration sensors, and a shaker was attached to the right wing to apply input signals. Nonlinearity in the structural dynamics, primarily originating from the mounting interfaces of the payloads, was investigated. The research focused on the back connection of the right-wing-to-payload interface, which was identified as a significant cause of nonlinear aberrations. Measurements were taken at a sample frequency of 400 Hz, with two different input signals: real force supplied by the shaker and voltage recorded at the signal generating amplifier's output. The output consisted of three acceleration signals measured at various locations on the payload and right wing (A. U. Jan, 2020).

2.5 Analysis

Detailed examination of how the data is used and applied to the aircraft maintenance to be able to find Any information required to guarantee that an aircraft or an aircraft component can be maintained in a way that ensures the aircraft's airworthiness, the component's serviceability, and the operational and emergency equipment, as applicable, is referred to as maintenance data. Analysis will be done on Matlab with Neural Fitting applying Levenberg-Marquardt algorithm. Applying different percentages of training data and layer number to find best result and then comparing with other people’s results.

Chapter 3

LITERATURE REVIEW

3.1 Machine Learning Implementation in Aircraft Maintenance Systems

3.1.1 Predicting the Remaining Useful Lifetime of Aircraft Components

The replacement of aircraft components is one of the most significant tasks performed by the aviation maintenance staff (Azevedo et al., 2019b). Instead of doing scheduled maintenance at predetermined intervals, replacement may be necessary based on the state of the aircraft component to maximize the lifetime of the aircraft equipment [(T. Wang et al., 2008), (Capodieci et al., 2020)]. An intelligent analysis of the data collected from sensors can be used to gain information on the state of various aircraft components (Azevedo et al., 2019a). Furthermore, machine learning methods use data to train and apply models about the degradation nature of various components to real-world conditions so as to forecast the remaining useful life of aircraft components (Mathew et al., 2017). The goal of this type of projects are to create a web interface for simulating a Prognostics and Health Management (PHM) system, in which multiple machine learning approaches can be utilized to forecast the Remaining Useful Life (RUL) of certain aircraft components (El Afia & Sarhani, 2017). Different approaches based on Artificial Intelligence are suggested and target accurate prediction of RUL of aircraft components (Rengasamy et al., 2018), (Seale et al., 2019). Models are trained with training dataset and validated with test dataset to achieve this (Ruan et al., 2021), (Doğru et al., 2020).

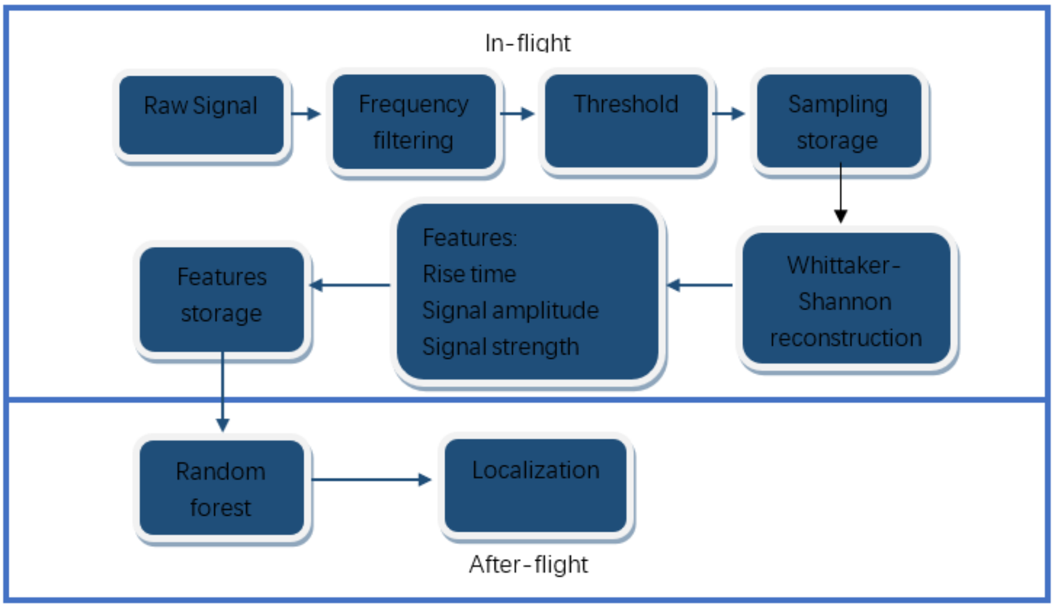


Figure 4: Schematic of Random Forest (Ai et al., 2021)

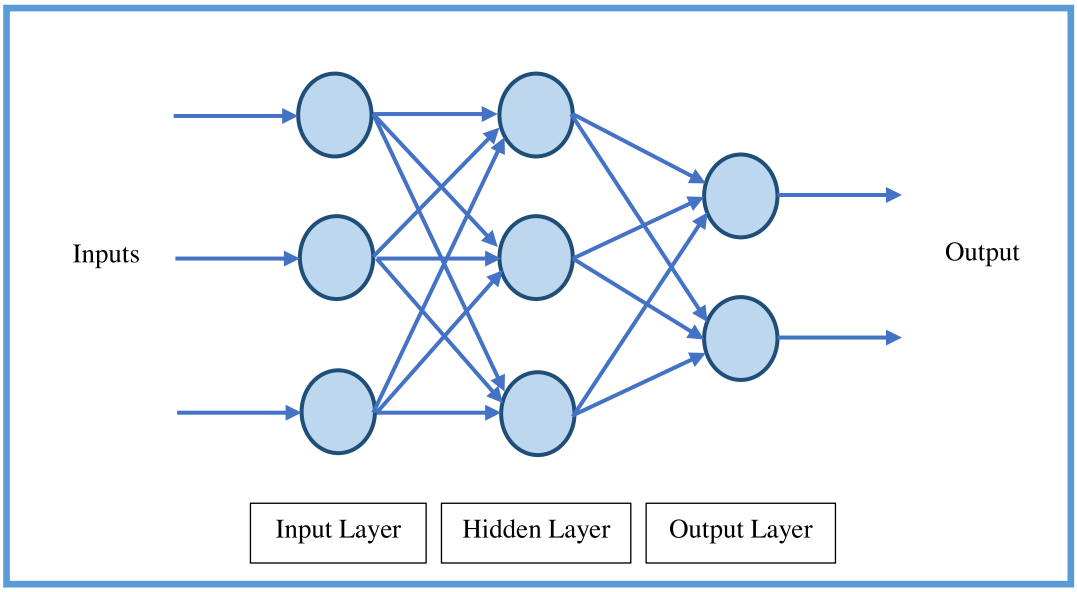


Figure 5: Schematic of a Neural Network (A. U. Jan, 2020)

3.1.2 Predictive Modelling of Aircraft Systems Failure

The aviation sector has a lot of information and maintenance data that might be utilized to get useful results in projecting future actions. A study shows the introduction of machine learning models that use feature extraction and data removal to anticipate failures in aviation system (W. Yan & Zhou, 2017). Corrective maintenance is performed by maintenance specialists on the ground using real-time condition monitoring data collected during the flight in recent aircraft maintenance systems. By analysing historical data from airplane maintenance systems, a predictive model is provided for predicting high-priority defects in advance, and preventive maintenance can be performed based on the model's prediction results. A study shows how the prediction of faults with varying priorities is treated as a binary classification problem. For decades, maintenance technology has advanced, and numerous ways to reducing the likelihood of machine system failure have been found. Corrective and preventive maintenance are the two types of maintenance approaches in general (Celikmih et al., 2020).

3.2 Objectives in implementing ML to Aircraft Maintenance

3.2.1 Replacement of Components of the Aircraft

To maximize equipment lifetime, replacement may be based on the condition of the aircraft components rather than a set time period. An intelligent analysis of sensor data can determine the condition of various aircraft components. In order to forecast the remaining useful life of aircraft components, machine learning techniques can be used to develop models trained on data and apply these models in describing the degradation behaviour of individual components to real-world settings (Azevedo et al., 2019b).

3.2.2 Maintenance of Parts

Artificial intelligence-based algorithms were used to define engine maintenance work, improve customer engine availability and maintenance costs, and obtain a procurement plan of integrated parts with intervention planning and maintenance strategy implementation. Machine Learning was used on a workshop dataset to optimize the amount, cost, and lead-time of warehouse spare parts in order to achieve this goal. This dataset comprises information such as repair claims, engine operating hours, forensic evidence, and general information regarding processed spare parts for a given engine type, spanning various years and fleets. These data have been used to develop various machine learning models in order to anticipate the repair state of each spare item for better warehouse handling. A multi-label classification strategy was utilized to create and train a Machine Learning model for each spare part that predicts the part repair state in the same way that a multiclass classifier does. Each classifier is specifically requested to predict the part's repair state (categorized as "Efficient," "Repaired," or "Replaced") (Capodieci et al., 2020).

3.2.3 Maintenance Planning

The process of developing a course of action is known as maintenance planning. Effective maintenance planning entails devising a strategy that encompasses all maintenance, repair, and construction tasks. Work to be done should be specified as clearly and completely as feasible (El Afia & Sarhani, 2017). In aviation, aircraft component reliability and availability have always been essential considerations. The reliability of aircraft components and systems will be improved through accurate prediction of possible faults. The entire maintenance and overhaul expenses of airplane components are determined by the maintenance operations schedule (Celikmih et al., 2020). In the aviation business, intelligent maintenance, repair, and overhaul (MRO) has become increasingly crucial. Aircraft are increasingly equipped with sensors that continuously collect data on their status, diagnosis, and potential defects. Effective maintenance management is aided by the ability to use sensor data to accurately predict and diagnose problems. Furthermore, the extensive use of sensors in aircraft has facilitated the move from time-based maintenance, in which maintenance is scheduled at regular intervals, to condition-based maintenance, in which choices are made based on data gathered through sensor monitoring (Rengasamy et al., 2018). Aircraft planning and scheduling is one of the most time-sensitive and important tasks in the airline business. Each flight leg on each aircraft will be covered exactly once during sequencing of the flight legs. This practice can be referred to as aircraft maintenance routing because it takes maintenance constraints into account while establishing each aircraft's routes because each aircraft must undergo routine maintenance inspections (Ruan et al., 2021).

3.2.4 Safety

Aircraft maintenance is a crucial task in the field of aeronautics. An airplane is sensitive to the effects of component failures due to component degradation over time, which could jeopardize the aircraft's overall reliability and safety. Aircraft maintenance is carried out by qualified professionals to prevent issues. The current practice of repositioning airplane parts at predetermined intervals complies with regulatory standards, however it does not maximize the useful lifetime of these components. Aviation maintenance's goal is to ensure the safety of the aircraft and its occupants by preventing aircraft subsystems from becoming vulnerable to failure and, if at all possible, extending component useful lifetimes to reduce costs for the airline industry (Azevedo et al., 2019a). According to Airbus system/component failure or malfunction (SCF) was responsible for roughly 13% of total losses between 1998 and 2017 (Airbus, 2024).

3.2.5 Maintenance Cost

In aviation, aircraft component reliability and availability have always been essential considerations. The reliability of aircraft components and systems will be improved through accurate prediction of possible faults. The entire maintenance and overhaul expenses of airplane components are determined by the maintenance operations schedule. Maintenance expenditures account for a large amount of an aviation system's total operating expenses (Celikmih et al., 2020).

Aviation Production Planning organizes aircraft maintenance plans, which include how to schedule maintenance and repairs at the right time, in the right location, and in the right order to guarantee that the actual flight plan is followed and that maintenance expenses are kept to a minimum. The airline's profitability has been significantly impacted by rising operating costs as a result of greater competition in aviation sector. Lowering operational expenses has become an important approach for civil aviation companies to reach new profit levels. Maintenance cost differ depending on the type of aircraft. This cost can vary from 10% to 45% of annual operational costs. Even if it's 10% that might not seem like a lot, when overall running costs could be in the thousands or millions of dollars (Brandon Battles, 2003). So, a main problem is how to minimize cost of maintenance and the upkeep time of an aircraft. Mathematical formulation as well as practical methodology are implemented to solve the problem of airplane maintenance scheduling. The purpose of the formulation and solution approach is to designate aircraft in order to reduce maintenance costs and time.

Table 2: Typical Objectives in Aircraft Maintenance

Objective	References	Remarks
Replacement of Components of the Aircraft	[(Azevedo et al., 2019b), (Azevedo et al., 2019a)]	• To extend the life of aircraft equipment, replacement may be necessary depending on the status of the aircraft components.
Maintenance of Parts	[(Capodieci et al., 2020), (Fernandes et al., 2020), (Godoy et al., 2016)]	• To optimize the amount, cost, and lead-time
Maintenance Planning	[(Celikmih et al., 2020), (Rengasamy et al., 2018), (Ruan et al., 2021)]	• Time-sensitive and crucial functions in the airline industry • Increase the asset's lifespan. The most crucial benefit of preventive maintenance, in my opinion, is that assets have a longer lifespan. • Breakdowns are less likely. • Boost your productivity. • Reduce unscheduled downtime.
Safety	[(Azevedo et al., 2019b), (Azevedo et al., 2019a), (Airbus, 2024)]	• An improved safety culture. • In the operational environment, there is a higher level of safety. • Higher levels of compliance. • There are fewer accidents. • There will be fewer inexcusable safety mishaps.
Maintenance Costs	[(Celikmih et al., 2020), (Brandon Battles, 2003)]	• Routine Maintenance Should Be Optimized • Reduce the amount of non-routine maintenance. • Improve the Aircraft Components' Reliability.

3.3 System Variables/Parameters Considered While Implementing ML in Aircraft Maintenance

3.3.1 Time for Repair

Predictive maintenance involves predicting maintenance needs in advance using time-based data from in-service assets such as trains and planes (Rengasamy et al., 2018). One main objective of this method is to be able to accurately forecast when it is time to repair or replace a component (El Afia & Sarhani, 2017). The advantages of prolonged usage are lost if done too far in advance; if done too late, unforeseen failures can occur, lowering asset availability (Ai et al., 2021). As a result, improving component lifetime estimates accuracy is a continuous priority (Seale et al., 2019).

3.3.2 Time for Replacement

The replacement of aircraft components is one of the most significant tasks performed by the aviation maintenance staff (Brandon Battles, 2003). Instead of doing scheduled maintenance at predetermined intervals, the condition of aircraft components should determine if replacement can be done in order to maximize the lifetime of the aircraft equipment (El Afia & Sarhani, 2017), (Ai et al., 2021).

3.3.3 Maintenance Cost

The major purpose of this study is to design Artificial Intelligence-based iterative algorithms P Korvesis et al. (2017) in order to define the engine removal plan and its repair work, and also get an acquisition plan of integrated components with intervention planning and maintenance strategy execution (Mathew et al., 2017).

3.3.4 Safety

The purpose of aviation maintenance is to guarantee the safety of the plane and its occupants by avoiding aircraft subsystems from becoming vulnerable to failure and, if feasible, extending component useful lifetimes to reduce costs for the airline industry (W. Yan & Zhou, 2017), (Mathew et al., 2017).

Table 3: System Variables/Parameters in Aircraft Maintenance

Variables	Units	References	Remarks
Time for Repair	seconds	[(El Afia & Sarhani, 2017), (Ai et al., 2021), (Seale et al., 2019), (D. K. Bhattacharya, 2017)]	• Forecasting repair time. • Avoids failures. • Ensures that aircraft will perform safely. • Repair on time. • Optimize regular repair. • Enhance the Reliability of Aircraft Components.
Time for Replacement	seconds	[(El Afia & Sarhani, 2017), (Ai et al., 2021), (Seale et al., 2019), (D. K. Bhattacharya, 2017)]	• Forecasting replacement time. • Maximize the equipment utilization lifetime. • Using time efficiently. • Optimize regular replacement.
Maintenance Cost	Dollars	[(Mathew et al., 2017), (P Korvesis, 2017)]	• Maintenance cost of parts and aircraft. • To achieve competitive advantage over competitors. • Reduce Non-regular Maintenance. • Optimization of Maintenance

3.4 Constraints in Utilizing ML in Aircraft Maintenance

3.4.1 Data might be interpreted incorrectly.

Resulting in erroneous maintenance requests. Data Collection: The amount of data required is directly proportional to the difficulty of the problem to be addressed. Data sources for a company might be both proprietary and open, such as weather data, traffic data, and so on. Data controlled by a business can be quantitative (loan amount, customer retention rate), category (gender, colour, property kind), time stamped (how many things were purchased over a period of time), or even free text (emails, doctor's notes).

Data transformation is required, as well as the removal of missing values. When data is acquired from many sources, the formats vary and must be standardized in order for ML to understand them. This level includes a lot of feature engineering. It also necessitates the creation of linkages between various sources of data. For more meaningful patterns, sales performance can be separated into day, month, and year category values.

Data Training: The analytics may now begin, and selecting the appropriate data model is critical - various methods are used for different jobs. In the 80:20 percent rule, it's critical to divide the data into training and assessment sets.

Parameter Tuning: The model will be evaluated against the evaluation set, and parameters such as the number of training steps, learning rate, and so on will be fine-tuned.

3.4.2 Contextual information

Such as the age of the equipment or the weather, may be ignored by predictive analysis. Time series data are made up of data points that are dated at specific moments in time. This data is often gathered at regular intervals. You can compare data from week to week, month to month, year to year, or any other time-based metric you want by learning and implementing time series data. Numerical data is just a collection of numbers that aren't rooted in any particular time periods, whereas time series data has set beginning and ending points.

3.4.3 Proactive

Physical examination and equipment maintenance may be discouraged by predictive maintenance. The process of examining objective structural findings through observation and testing is known as physical examination. The information gathered must be carefully linked with the history of the parts. Furthermore, a thorough physical examination should offer 20% of the information needed for parts inspection and treatment. Rather than actual machine conditions, may trigger preventative maintenance operations. To make sure that any issues are minimized, frequent inspections, maintenance, and service are all carried out. Thus, only the most precise and reliable aircraft parts must be used, and there must be a concerted effort to minimize human error. During inspections, the quality of components and aircraft parts is evaluated using routine manual checks and visual inspections. The main objective is to maintain aircraft in excellent condition to prevent any type of failure that could cause an accident. "Regularly scheduled inspections and preventative maintenance maintain airworthiness," according to the Federal Aviation Administration (FAA). By identifying minor faults and wear and tear early on and enabling specialists to fix and rectify them swiftly, maintenance of aircraft components and services reduces malfunctions and the danger of operational failures. Keeping accurate records is also beneficial to the process.

Table 4: Typical Constraints in Aircraft Maintenance

Constraints	References	Remarks
Data	[(Mathew et al., 2017), (Ai et al., 2021), (Seale et al., 2019), (Brandon Battles, 2003), (P Korvesis, 2017)]	• Data Collection • Data Transformation • Data Training • Parameter Tuning
Daily Range	[(Mathew et al., 2017), (Ai et al., 2021), (Seale et al., 2019), (Brandon Battles, 2003), (P Korvesis, 2017)]	• Time Series Data
Physical Examination	[(Mathew et al., 2017), (Ai et al., 2021), (Seale et al., 2019), (Brandon Battles, 2003), (P Korvesis, 2017)]	• Part Examination • Physical Examination
Condition	[(Mathew et al., 2017), (Ai et al., 2021), (Seale et al., 2019), (Brandon Battles, 2003), (P Korvesis, 2017)]	• Risk of Failures • Avoid Breakdowns

3.5 Simulation Platforms used for Applications of ML in Aircraft Maintenance

Data Science and Machine Learning Platforms offer platforms for the development, implementation, and analysis of machine learning algorithms (Çinar et al., 2020). To do predictive maintenance, sensors are first installed in the system to monitor and gather data on its activities. Time series data is used in predictive maintenance. A timestamp, a collection of sensor readings recorded at the same time as timestamps, and device IDs are all included in the data. The purpose of predictive maintenance is to anticipate whether equipment will fail in the near future at time *t* using data collected up to that point. Two approaches to predictive maintenance are: (1) the classification technique which indicates if the following n-steps are likely to fail, (2) the regression technique which forecasts the amount of time until the next failure. This is referred to as Remaining Useful Life (RUL).

3.5.1 Matlab

MATLAB is a computer application used by engineers and scientists to study and build systems and products. The MATLAB language enables the most natural representation of computational mathematics (Azevedo et al., 2019b). You can create models and applications by analysing data, developing models and developing algorithms. MATLAB allows you to take your ideas from research to production by deploying them to enterprise applications and embedded devices and integrating them with Simulink and Model-Based Design (El Afia & Sarhani, 2017). MATLAB contains interactive and visualization tools that make machine learning operations simple. One may study your data, uncover essential attributes, and share your results using data visualization tools.

3.5.2 Python

Python is an interpreted high-level, general-purpose programming language. İts design philosophy uses of considerable indentation in order aid code readability. Its language features and object-oriented methodology are designed to assist programmers in writing logical code for both small and large-scale projects. Python supports a variety of programming standards, including structured (especially procedural) programming, object-oriented programming, and functional programming (Ruan et al., 2021). It is sometimes referred to as a "batteries included" language due to its extensive standard library (El Afia & Sarhani, 2017).

3.5.3 R Programming

R is a statistical computing and graphics programming language. It is usually used by statisticians and data miners to design statistical software and for data analysis. R is widely used, according to polls, data mining surveys, and examinations of scientific literature databases (Mathew et al., 2017). The syntax will come to you shortly. You don't have to be an expert coder. It's not about being a great coder; it's about understanding which packages to use and how to use them effectively. Using R programming is that easy.

3.5.4 ERP Platform

Enterprise resource planning (ERP) is a management and integration technique used by organizations for their daily operations (Celikmih et al., 2020). Numerous ERP software programs combines all of the tasks necessary to run a company into a single system. ERP software can integrate planning, purchasing, marketing, finance, sales, inventory, human resources, and other functions (Seale et al., 2019).

Table 5: Typical Simulation Platforms

Simulation platforms	References	Remarks
MATLAB	[(Azevedo et al., 2019b), (El Afia & Sarhani, 2017), (Azevedo et al., 2019a), (Paluszek & Thomas, 2019)]	• Enhance fast implementation and testing of algorithms. • Creating the computational codes is easy. • Debug with ease. • Big database of pre-installed algorithms. • It's simple to manipulate still photographs and make simulation films. • Symbolic computation is simple to do. • Add external libraries easily.
PYHTON	[(Mathew et al., 2017), (Singh et al., 2020), (Ruan et al., 2021), (Doğru et al., 2020), (Raschka et al., 2020)]	• It is simple to read, learn, and write. • Python is a high-level programming language • Syntax that is similar to English. • Increased productivity. • Language that has been translated. • Typed in a dynamic way. • It's free and open-source.
R PROGRAMMING	[(Singh et al., 2020), (Ruan et al., 2021)]	• Superb for Statistical Analysis. • Open-source. • Includes a variety of libraries • Support for several platforms. • Supports a variety of data types.
ERP PLATFORM	[(Doğru et al., 2020), (Seale et al., 2019)]	• By assisting users in navigating complicated procedures, ERP improves efficiency and productivity. • blocking re-entry of data • Production, order fulfilment, and delivery are all operations that may be improved. Throughout, procedures have been streamlined and made more efficient.

Chapter 4

METHODOLOGY

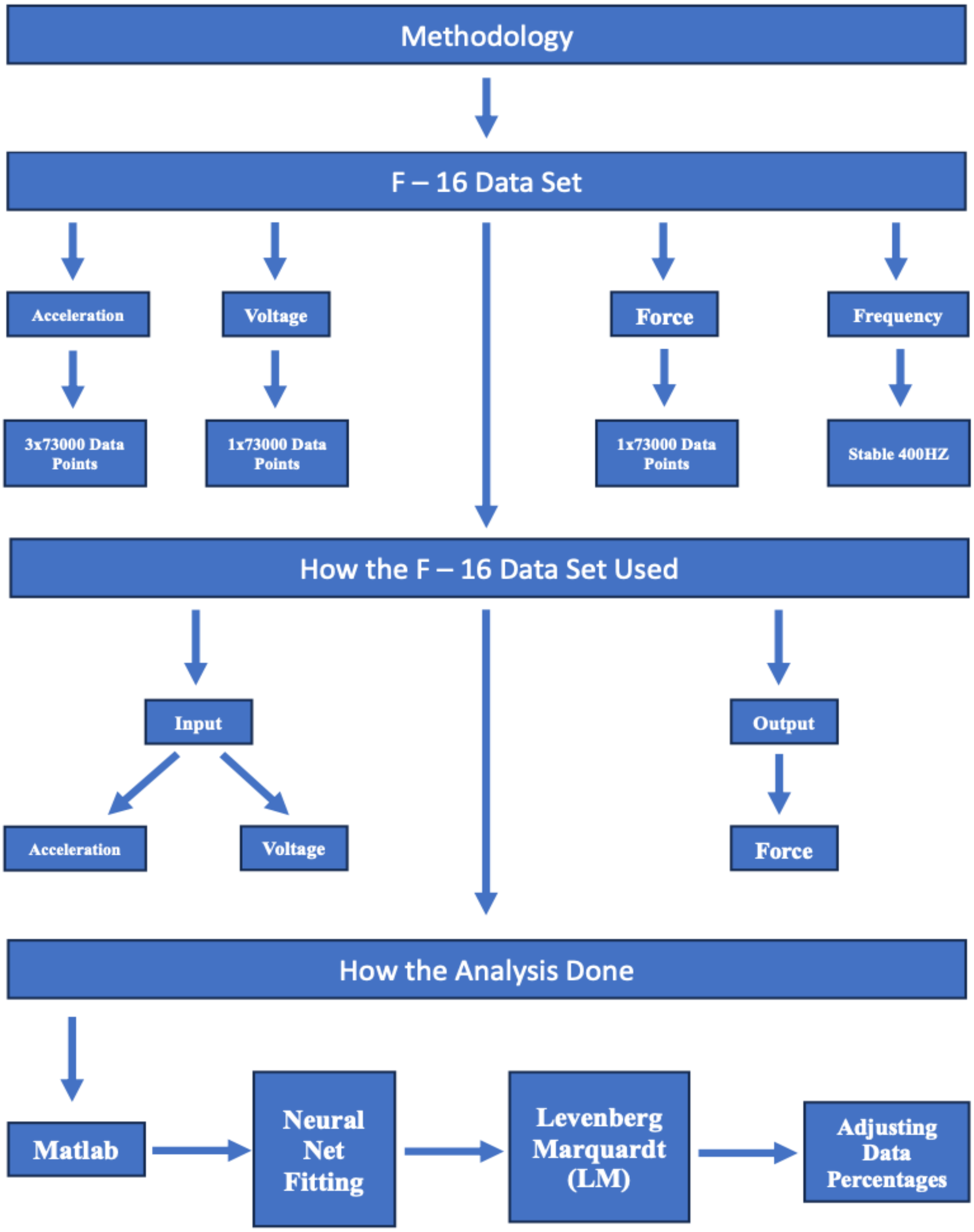


Figure 6: Flow Chart for Methodology

4.1 F-16 Data Set

The F-16 data includes parameters of Acceleration, Voltage, Force and Frequency (Constant). Each F-16 data set(combined 4 of them) includes around 73000 data points with 3 reparations for Acceleration and also, includes around 73000 data points with 1 reparation for Voltage and Force. A group of connected, discrete pieces of related data that may be viewed separately, together, or as a single entity is called a data set. A data structure of some kind is used to arrange a data set. F-16 data set is going to be used for this research. At the tips of the wings, two fake payloads were installed to replicate the mass and inertia characteristics of actual equipment that normally accompanies an F-16 during flight. One hundred and forty-five acceleration sensors were used to instrument the aircraft structure. Under the right wing, a single shaker was affixed to apply input signals. The mounting interfaces of the two payloads were predicted to be the primary source of nonlinearity in the structural dynamics. These interfaces are made up of T-shaped connection parts that slide through a rail that is fastened to the wing side of the payload. Focusing on the back connection of the right-wing-to-payload interface, which was shown to be the primary cause of nonlinear aberrations in the aircraft dynamics based on an initial investigation. 400 Hz was used as the sample frequency for the measurements. There are two different input signals available. The real force supplied by the shaker and measured by an impedance head at the excitation position, as well as the voltage recorded at the signal generating amplifier's output, which serves as a reference input. The output quantities are three acceleration signals. They were measured on the payload adjacent to the nonlinear interface of interest, at the excitation position, and on the right wing (Maarten Schoukens et al. 2017). On the left-hand side dummy payload mounted at the right-wing tip. On the right-hand side connection of the right-wing-to-payload mounting interface. Reproduced under CC BY-SA 4.0 from (Alessandrini et al., 2022; Perne et al., 2021)

Figure 7: Fake Payloads at the Places

4.1.1 Excitation Using Sine-Sweep and a Linearly Negative Rate

F16 Data Sine show that sine-sweep excitations were administered at a linear, negative rate of 0.05 Hz/s (sweep down). The input frequency range covered was between 15 and 2 Hz. As benchmark data, seven distinct excitation levels are given. One may classify the lowest level at 4.8 N input amplitude as a linear data collection. Three higher excitation levels data sets number 3, 5, and 7 which correspond to 28.8, 67.0, and 95.6 N, respectively are provided to serve as estimation data in nonlinear vibration regimes. The models estimated using the data sets 3, 5, and 7 are to be tested using datasets 2, 4, and 6 at 19.2, 57.6, and 86.0 N, respectively.

4.1.2 Full Frequency Grid Multisine Excitation

Data with a complete frequency grid spanning from 2 to 15 Hz, recorded under multisine excitations, are presented. Nine periods were recorded at each force level, taking into account a single input signal realization. There are 8192 points in a period. Keep in mind that the initial measurement period contains transients. Seven excitation levels are taken into consideration, beginning with linear data at 12.4 NRMS (data set 1), just like in the sine-sweep instance. Furthermore, matching test sets (numbers 2, 4, and 6 at 24.6, 61.4, and 85.7 NRMS, respectively) are provided for three nonlinear estimation data sets (numbers 3, 5, and 7 at 36.8, 73.6, and 97.8 NRMS, respectively).

4.1.3 Excitation of a Multisine Wave Using a Random Frequency Grid

Additionally, multisines were used while only odd frequencies were stimulated. Additionally, one frequency line was arbitrarily discarded from each set of four consecutively excited odd lines in order to serve as a detection line for odd nonlinearities. The frequency range of 1 to 60 Hz was excited in this configuration. Each level has three recorded periods, with each level taking into account ten input realizations. There are 16384 points in a period. Keep in mind that each realization is only in a constant state for the final two periods. 200 Hz was the initial sampling rate for the datasets. Assuming the data is periodic and in a steady state, they were processed period by period and up sampled to 400 Hz in the frequency domain. The right wing is excited by a shaker; the voltage (reference) and force (input) signals of the shaker are measured. The nonlinearities primarily stem from the payload connection interface where acceleration is measured. The right wing is excited by a shaker using odd multisines, with one random bin skipped within each group of four successively excited odd lines.

Figure 8: F-16 Ground Vibration Testing Measurement Campaign

4.2 Usage of F-16 Data Set

The data used is related to the paper that is released in 2017 (Maarten Schoukens, 2017). Since the aviation is a restricted research area for many countries, it is not easy to reach every data and information (They can be used against that country who is trying to defence or enhance their technology) for this kind of research. That's the first point that F-16 data is chosen and the other point is that, data size is huge so the results will be satisfying. Two separate input signals are provided: the voltage at the output of the signal generator amplifier, which serves as a reference input, and the actual force from the shaker measured by an impedance head at the excitation location. The output consists of three acceleration signals. Data is minimized for correct and best usage and also, analysed using different algorithm and model. The recent research of the papers that used F-16 data with different models are given below.

Table 6: Different Kind of Models That Used F-16 Data

Model	References	Remarks
SINDy	(NINNI, 2023)	• Sparse Identification of Nonlinear Dynamics • A symbolic sparse regression problem, to identify nonlinear systems. • Autoencoder network. • An artificial neural network called an autoencoder is utilized for unsupervised learning, or the efficient coding of unlabeled input.
FOURIER NEURAL OPERATOR	(Park et al., 2023)	• Capable of managing time continuously without the need for iterative processes. • A type of deep learning systems with the purpose of learning function space maps that are infinite-dimensional. • The first machine learning technique to effectively simulate turbulent flows at zero-shot super-resolution.
NODE	(Gong et al., 2023)	• Neural Ordinary Differential Equation. • A deep learning model integrated with differential equations. • A viable method for nonlinear system identification assignments. • A neural network architectural style that blends deep learning and ordinary differential equations (ODEs) ideas.
BLA	[(Csurcsia, Decuyper, Runacres, et al., 2023),(Csurcsia, Decuyper, Renczes, et al., 2023)]	• Best Linear Approximation. • An important tool for initiating the nonlinear modeling process and analyzing nonlinear systems. • Framework of MIMO (multiple input, multiple output) systems.

		• Large multivariate polynomials found in PNLSS models are reduced to a simplified representation through function decoupling.
SAMI	(Csurcsia et al., 2020)	• Simplified Analysis for Multiple Input Systems. • Designed for the industrial measurement of multiple-input vibro-acoustic systems. • For easily analyzing observations of multiple-input multiple-output (MIMO) nonlinear systems. • Methods include BLA, LPM (local polynomial method), LRM (local rational method),
NONLINEAR ARX	(Wigren & Schoukens, 2013)	• Autoregressive with Extra Input • The linear ARX models are extended to the nonlinear case by nonlinear ARX models. • A combination of an AR model and an X model. • Least squares method.
PARTICLE-BERNSTEIN POLYNOMIALS	(Alessandrini et al., 2022)	• Nonlinear Dynamic System Identification in the Spectral Domain. • Regression using particle Bernstein polynomials in machine learning for nonlinear system identification
GAUSSIAN PROCESS MODELLING	(Perne et al., 2021)	• A scalable multi-output general purpose model (GP) that can easily and tractably describe nonlinear, potentially input-varying, dependencies between outputs. • A nonparametric supervised learning technique called Gaussian Processes (GP) is used to resolve probabilistic classification and regression issues. • A stochastic process, which is typically an ensemble of random variables that are indexed spatially or temporally. • A nonparametric supervised learning technique for probabilistic classification and regression issues.
DEEP CONVOLUTIONAL NETWORK	(Maarten Schoukens, 2017)	• The kind most frequently employed to find patterns in pictures and videos. • A deep learning network design that uses data to inform its learning. • The amount of layers in the network is what makes a "Deep CNN" "deep."

4.3 How the Analysis Done with ANN

By using MATLAB ANN model (Neural Net Fitting from Application) with decent layer number. A competent network design can be programmatically constructed, and the number of hidden layers and nodes in an ANN can be effectively determined by adhering to a modest set of explicit principles.

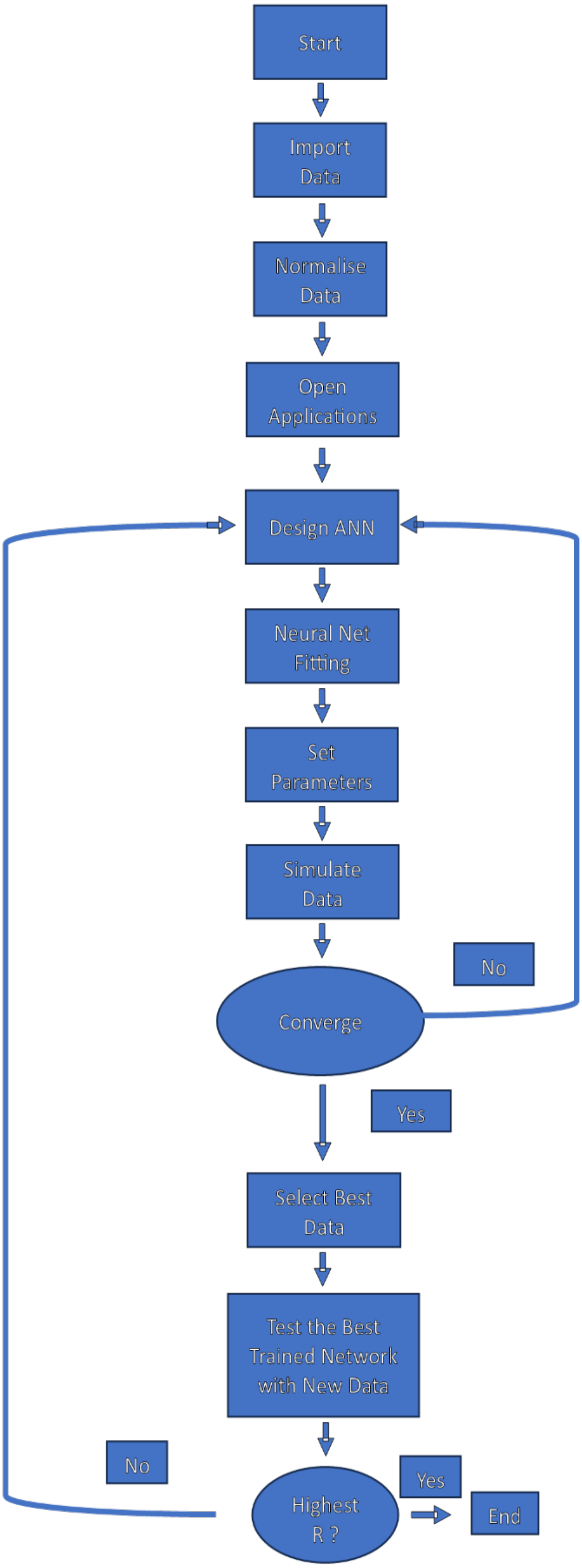


Figure 9: ANN Flowchart

For choosing hidden layers, you won't need any hidden layers at all if the data is linearly separable. Neural networks with one to two hidden layers would function well with data that is less complicated and has fewer dimensions or attributes. Three to five hidden layers can be employed to provide the best result when dealing with data that has several dimensions or features. Adding more hidden layers would also make the model more complex, and selecting hidden layers with high numbers or in two digits could occasionally cause overfitting. For selecting nodes in concealed layers, the number of nodes in each hidden layer must be chosen after hidden layers have been determined. The size of the input layer and the output layer should be in between the number of hidden neurons. The sqrt (input layer nodes * output layer nodes) number of hidden neurons is the most suitable quantity. In order to identify the target class and grow closer to pattern and feature extraction, the number of hidden neurons in each succeeding layer should keep decreasing. The algorithms mentioned above are merely examples of broad use cases; specific use cases might be tailored to them. For an Engine data set, despite being a function approximation problem, the Levenberg-Marquardt (LM) algorithm's superiority is not as evident. In this instance, the network has a far larger number of parameters, and the LM algorithm's benefits diminish as the number of parameters rises (MathWorks, 2024). Different percentage number of Training Data, Validation Data, Test Data and Layer Size is used and compared.(Karaoğlu et al., 2022)

4.4 How the Analysis Done with Linear Regression

Regression analysis is a statistical method used to estimate the relationships between multiple variables. The variable you aim to understand and predict is called the dependent or criteria variable. The variables that might influence the dependent variable are known as independent variables, also called explanatory variables or predictors. This analysis helps to quantitatively determine which independent variables affect the dependent variable and how changes in the independent variables impact the dependent variable. A regression analysis model relies on the sum of squares to mathematically determine how data points are distributed. The goal of the model is to find the smallest sum of squares and draw a line that best fits the data. In statistics, there's a distinction between simple and multiple linear regression. Simple linear regression uses a linear function to model the relationship between one independent variable and one dependent variable. In contrast, multiple linear regression involves two or more explanatory variables to predict the dependent variable. If the data shows non-linear correlations, nonlinear regression should be used instead of linear regression to model the dependent variable as a non-linear function. Step by step for linear regression is followed by; First, go to file > options in excel. Secondly, in the excel options dialog box, click on add-ins from the left menu. ensure that excel add-ins is selected in the manage box, then click ok. Then in the add-ins dialog box, check the box for analysis toolpak and click ok. And then, on the data tab, find the analysis group and click on the data analysis button. Select regression and click ok. Followed by the regression dialog box, configure the following settings: for the dependent variable, set the input y range. for the independent variable, set the input x range. To create a multiple regression model, select two or more adjacent columns that contain different independent variables. If your X and Y ranges have headers, make sure to check the Labels box. Choose where you want the output to appear; in this example, we select a new worksheet. To see the difference between the actual and predicted values, check the Residuals option. Click ok to see the regression analysis results generated by Excel. In our case dependent variable Force and Voltage and the others are independent variables are Acceleration 1/2/3 and Frequency.

Multiple R is a measure of the strength of a linear relationship between two variables, with values ranging from -1 to 1 indicating the strength of the relationship. An absolute value closer to 1 indicates a stronger relationship, with 1 representing a strong positive relationship, -1 representing a strong negative relationship, and 0 indicating no relationship.

R Square, or the Coefficient of Determination, is used to assess the goodness of fit of a regression analysis model, showing how well the points fit the regression line. An R2 value of 0.91 in our example indicates that 91% of the values fit the regression model, with 91% of the dependent variables explained by the independent variables. An R Squared value of 95% or higher is generally considered a good fit.

Adjusted R Square adjusts the R Square value for the number of independent variables in the model, making it more suitable for multiple regression analysis.

Standard Error is a measure of the precision of a regression analysis, with smaller values indicating a more accurate regression equation. While R2 explains the percentage of the dependent variable's variance explained by the model, Standard Error provides an absolute measure of the average distance of data points from the regression line.

4.5 How the Analysis Done with Regression Tree (Random Forest)

A regression tree ensemble is a model made up of multiple weighted regression trees that can improve predictive performance. LSBoost can be used to enhance regression trees, while fitrensemble or TreeBagger can be used to bag regression trees or create random forests. To implement quantile regression with a bag of regression trees, use

TreeBagger. For classification ensembles, such as boosted or bagged classification trees, random subspace ensembles, or ECOC models, refer to Classification Ensembles. Regression models can be trained using the Regression Learner app, where you can explore, select features, define validation schemes, train models, and evaluate outcomes. Automated training can help find the optimal regression model type for support vector machines, neural networks, linear regression models, regression trees, and Gaussian process regression models. In supervised machine learning, a model is created using known observations of predictors and responses to predict reactions to new input data. The trained model can be exported for future use or to gain insights. Training a model in Regression Learner involves two steps: validated model and complete model. The validated model uses a validation strategy to prevent overfitting, while the complete model is trained with all available data. The app displays the validated model results, including diagnostic metrics and graphs. Once one or more models are automatically trained, you can compare validation results and select the best fit for your regression problem. The model can be exported for forecasting fresh data without any delay.

Manuel Regression Model Training involves selecting a model type on the Learn tab. You can train models individually or as a group by clicking on the model type and then selecting the Train option. Repeat the process to explore different models or train all models at once by selecting the Train All option. In Automated Regression Model Training with Regression Learner allows users to automatically train multiple regression models on their data. Users can quickly try different models simultaneously and explore promising models interactively. If users have a specific model in mind, they can opt for manual regression model training instead. To start, open the

Regression Learner app, select data, choose predictors and response variable, train models, and explore the results. For further customization and optimization, users can also try non-optimizable model presets.

Chapter 5

RESULTS

5.1 Analysis Results ANN

The following table compares the R – Value results for each iteration of ANN analysis. According to the numerical ID for example; 60-20-20-5 means 60 percent of the Training Data, 20 percent of Validation Data, 20 percent of Test Data and 5 Layers are used.

Table 7: Best R-Value Result

Serial Number	Rtrain	Rvalidation	Rtest	Rall
60-20-20-5	0,92592	0,92637	0,92581	0,92599
60-20-20-10	0,92676	0,92684	0,92639	0,92671
70-10-20-5	0,92605	0,92594	0,92582	0,92599
70-10-20-10	0,92635	0,92575	0,92656	0,92633
70-20-10-5	0,92598	0,92633	0,92616	0,92606
70-20-10-10	0,92672	0,92699	0,92563	0,92667
75-10-15-5	0,92615	0,92613	0,92591	0,92611
75-10-15-10	0,92677	0,92646	0,9261	0,92664
75-15-10-5	0,92575	0,92706	0,92622	0,926
75-15-10-10	0,92660	0,92633	0,92645	0,92654
80-10-10-5	0,92589	0,92682	0,92623	0,92503
80-10-10-10	0,92675	0,92717	0,92729	0,92684

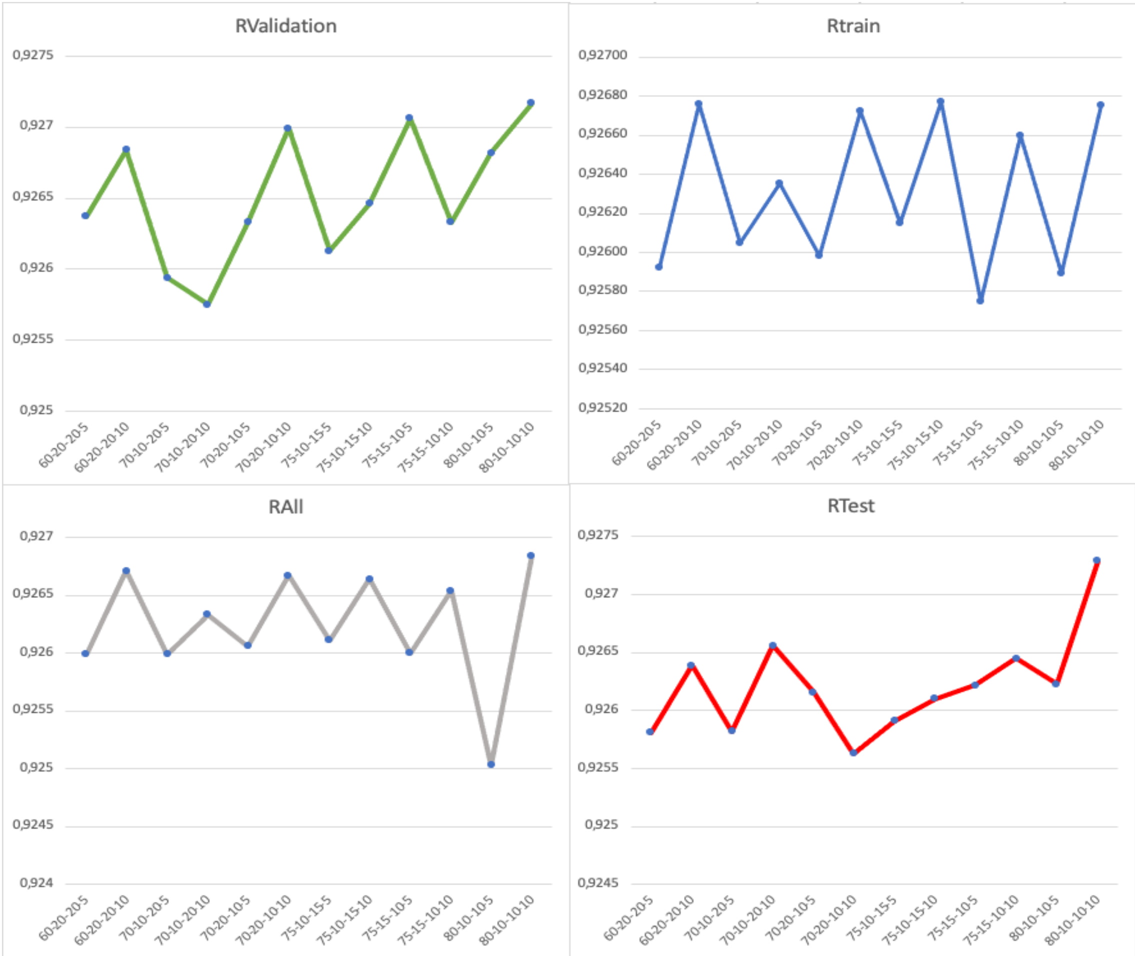


Figure 10: R-Value Results

According to the analysis results best R values were found at the analysis which contains 10 number of layers also we can clearly see that the best results are the results which we have used 80 percentage of data as training data. 80 percent of data is used as training data, 10 percent of data is used for test and validation data with 10 number of layers.

Table 8: Best MSE Results

Serial Number	MSEtrain	MSEvalidation	MSEtest	MSEall
60-20-20-5	3,3835E-05	3,3321E-05	3,4062E-05	3,3739E-05
60-20-20-10	3,3414E-05	3,3198E-05	3,3262E-05	3,3291E-05
70-10-20-5	3,3721E-05	3,3517E-05	3,4055E-05	3,3764E-05
70-10-20-10	3,3414E-05	3,3198E-05	3,3262E-05	3,3291E-05
70-20-10-5	3,3586E-05	3,3634E-05	3,4453E-05	3,3891E-05
70-20-10-10	3,2874E-05	3,2892E-05	3,3565E-05	3,3110E-05
75-10-15-5	3,3509E-05	3,3381E-05	3,4385E-05	3,3758E-05
75-10-15-10	3,2866E-05	3,3630E-05	3,3138E-05	3,3211E-05
75-15-10-5	3,3888E-05	3,3456E-05	3,3263E-05	3,3536E-05
75-15-10-10	3,3143E-05	3,3107E-05	3,2740E-05	3,2997E-05
80-10-10-5	3,3900E-05	3,3217E-05	3,2953E-05	3,3357E-05
80-10-10-10	***3,2788E-05***	***3,2572E-05***	***3,2448E-05***	***3,2603E-05***

The average of the squares of the mistakes, or the average squared difference between the estimated values and the actual value, is measured by the mean squared error (MSE) of an estimator (a process for estimating an unobserved variable). As the expected value of the squared error loss, MSE is a risk function. Because of randomness or the estimator's failure to take into consideration data that could lead to a more accurate estimate, the mean square error (MSE) is nearly always strictly positive (rather than zero). When discussing empirical risk reduction in machine learning, mean square error (MSE) can be used to estimate the genuine MSE, which is the average loss on the actual population distribution, by comparing it to the empirical risk, which is the average loss on an observed data set. So that lower MSE result gives higher accuracy for model. Again, according to the results best results are the model which is 80 percent of data is used as training data, 10 percent of data is used for test and validation data with 10 number of layers. So that this model and result is going to be compared with other papers results.

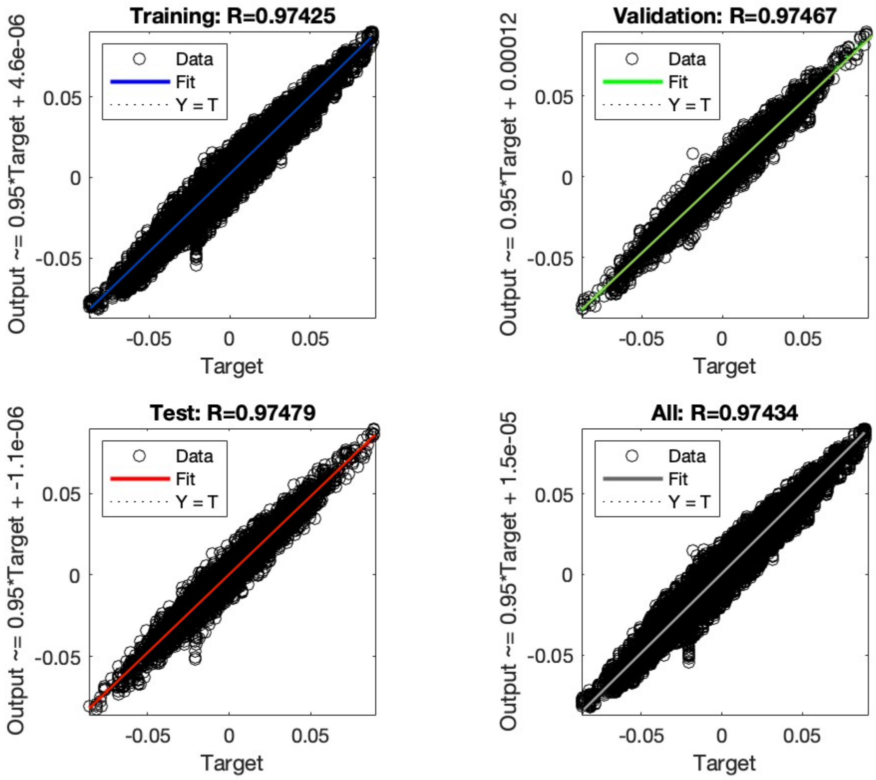


Figure 11: Best ANN Result

The validation loss was monitored to assess the model's effectiveness, and mean-absolute error (MAE) proved to be the most successful loss function, closely followed by mean-squared error (MSE). The smallest MAE obtained with 80-10-10 loss function. The statistics verify that the mistakes have the same variance and are, on average, of the same order of magnitude. In order to lower the maximum magnitude error, MSE can be used instead. It is doubtful that the MSE's replacement will have any further negative effects because it almost exactly matched the MAE's overall performance. Figure displays the 4G accelerometer components for the full test set. Each dot in the figure represents a network trial, and the units are G's. The places where the wing is closest to zero acceleration in each scenario exhibit the highest variance in the results (keep in mind that Earth's gravity is included in the zero acceleration in the vertical direction). The prediction error variances narrow as the wing oscillation/acceleration moves away from zero. There are a few obvious situations of significant inaccuracy for both longitudinal and lateral acceleration; these are probably the result of the MAE cost function's light penalty of greater errors. Compared to the other two directions, the vertical acceleration data points cover a far wider area.

Table 9: Predicted vs Actual Acceleration

Predicted Acceleration			Actual Acceleration		
Serie 1			Serie 2		
Acceleration 1	Acceleration 2	Acceleration 3	Acceleration 1	Acceleration 2	Acceleration 3
0,2501	0,2079	0,2123	-0,0184	-0,0496	-0,0463
0,2501	0,2300	0,2378	-0,0184	-0,0332	-0,0275
0,2534	0,2486	0,2580	-0,0160	-0,0195	-0,0126
0,2465	0,2694	0,2920	-0,0211	-0,0041	0,0126
0,2542	0,2905	0,3093	-0,0154	0,0115	0,0254
0,2573	0,3117	0,3294	-0,0131	0,0271	0,0402
0,2553	0,3375	0,3640	-0,0145	0,0462	0,0658
0,2720	0,3625	0,3864	-0,0022	0,0647	0,0824
0,2797	0,3907	0,4185	0,0035	0,0855	0,1061
0,2934	0,4253	0,4510	0,0136	0,1111	0,1301
0,3159	0,4587	0,4801	0,0302	0,1358	0,1517
0,3250	0,4986	0,5180	0,0370	0,1653	0,1797
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
0,1663	-0,0342	-0,0109	-0,0804	-0,2286	-0,2114
0,1867	0,0046	0,0234	-0,0653	-0,1999	-0,1860
0,2079	0,0421	0,0485	-0,0496	-0,1722	-0,1675
0,2194	0,0799	0,0867	-0,0411	-0,1442	-0,1392
0,2388	0,1187	0,1184	-0,0267	-0,1156	-0,1158
0,2575	0,1507	0,1468	-0,0129	-0,0919	-0,0948
0,2648	0,1835	0,1808	-0,0075	-0,0677	-0,0697
0,2796	0,2129	0,2044	0,0034	-0,0459	-0,0522
0,2867	0,2358	0,2293	0,0086	-0,0290	-0,0338
0,2903	0,2611	0,2554	0,0113	-0,0103	-0,0145
0,2966	0,2777	0,2718	0,0160	0,0020	-0,0024

Figure 12: Predicted vs Actual Acceleration 1

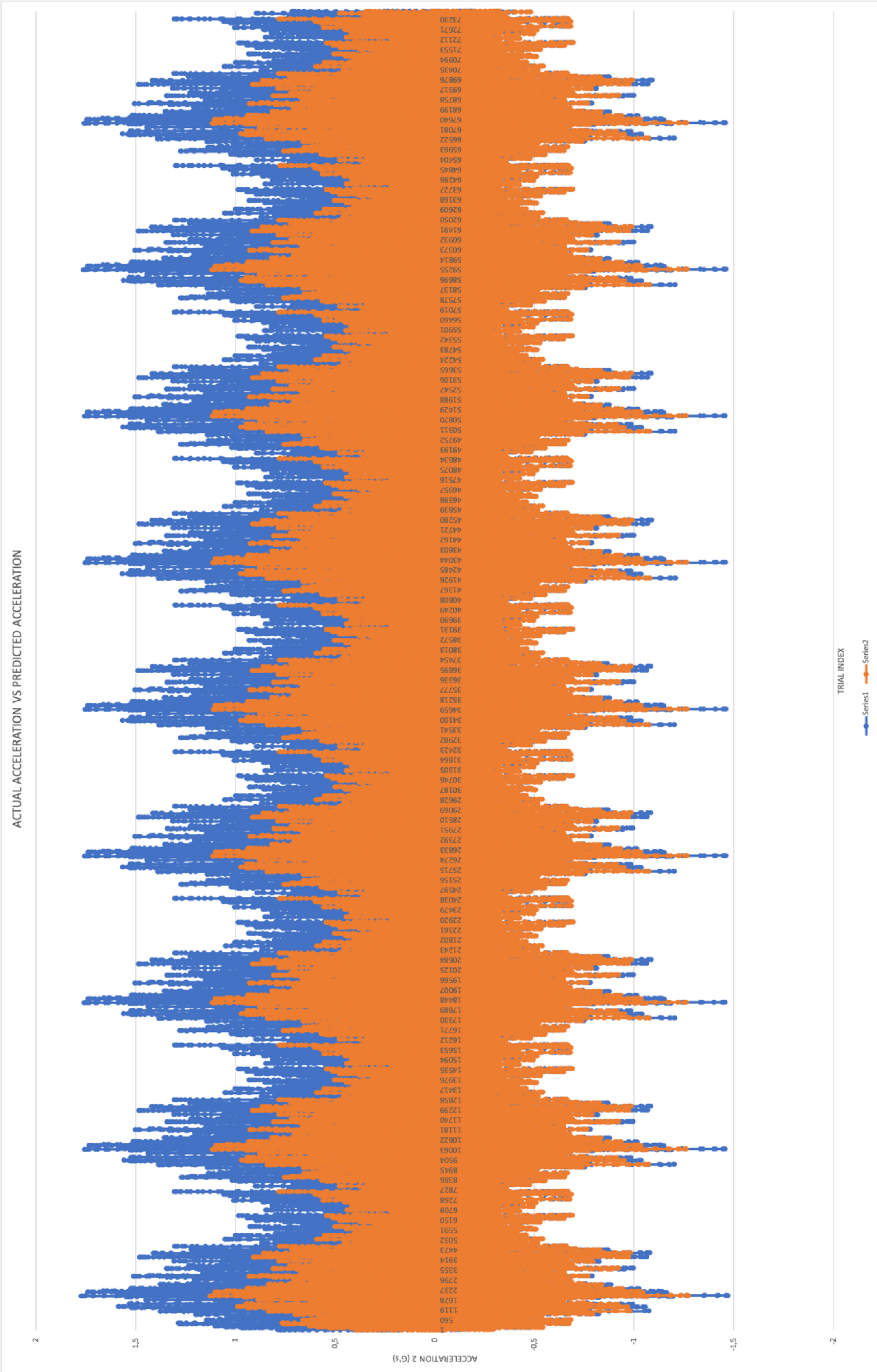


Figure 13: Predicted vs Actual Acceleration 2

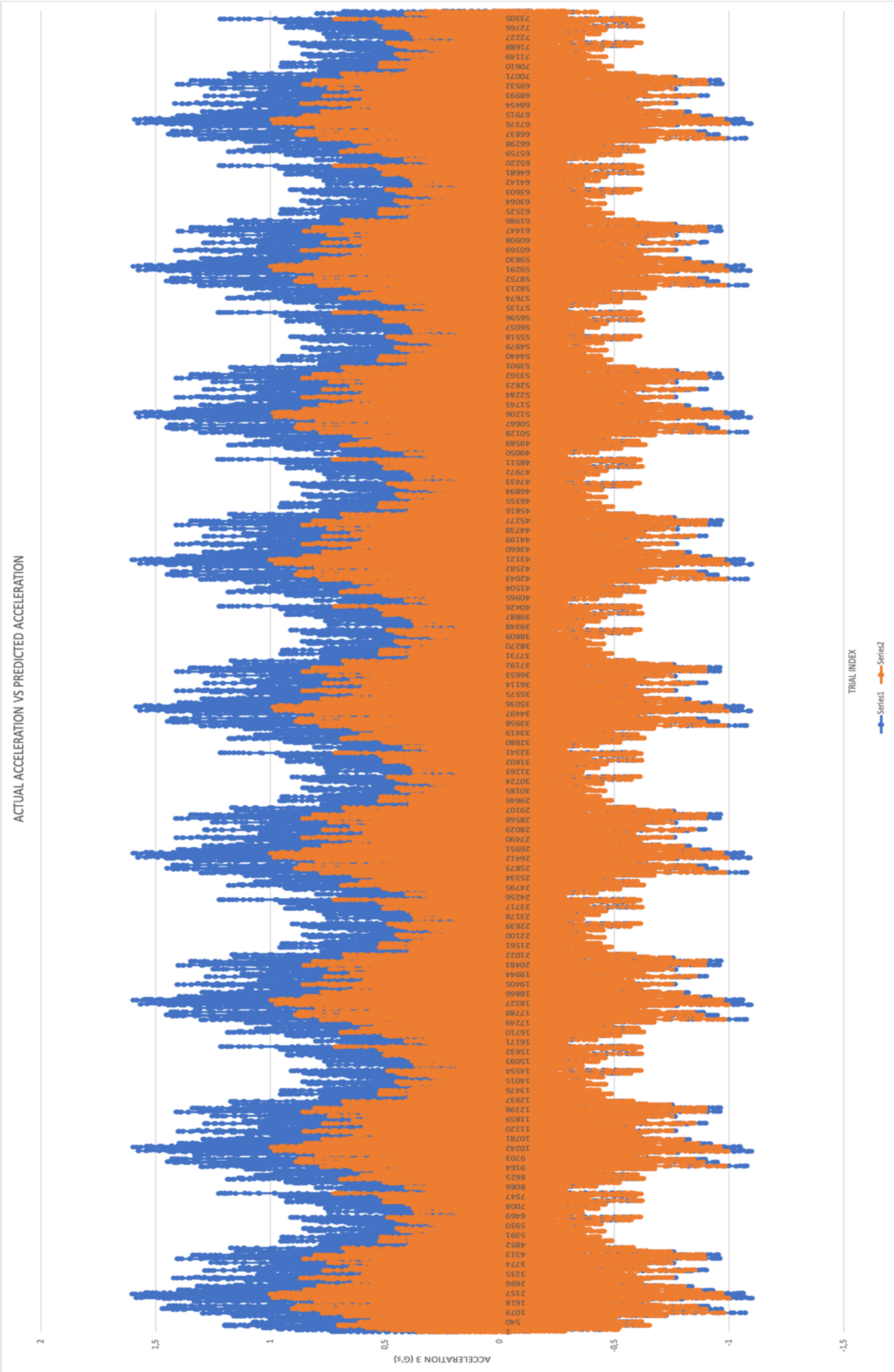


Figure 14: Predicted vs Actual Acceleration 3

5.2 Analysis Results Linear Regression

These regression results indicate a strong relationship between the variables, with an R square value of 0.8068 suggesting that 80.68% of the variability in the dependent variable can be explained by the independent variables. The adjusted R square value of 0.8055 accounts for the number of predictors in the model, and a multiple R value of 0.8982 indicates a high level of correlation. The low standard error suggests that the model has good predictive accuracy. Overall, these results suggest that the regression model is a good fit for the data.

Table 10: Linear Regression Results

Linear Regression Results	
Multiple R	0,8982
R Square	0,8068
Adjusted R Square	0,8055
MSE	4,26E-05

5.3 Analysis Results Random Forest

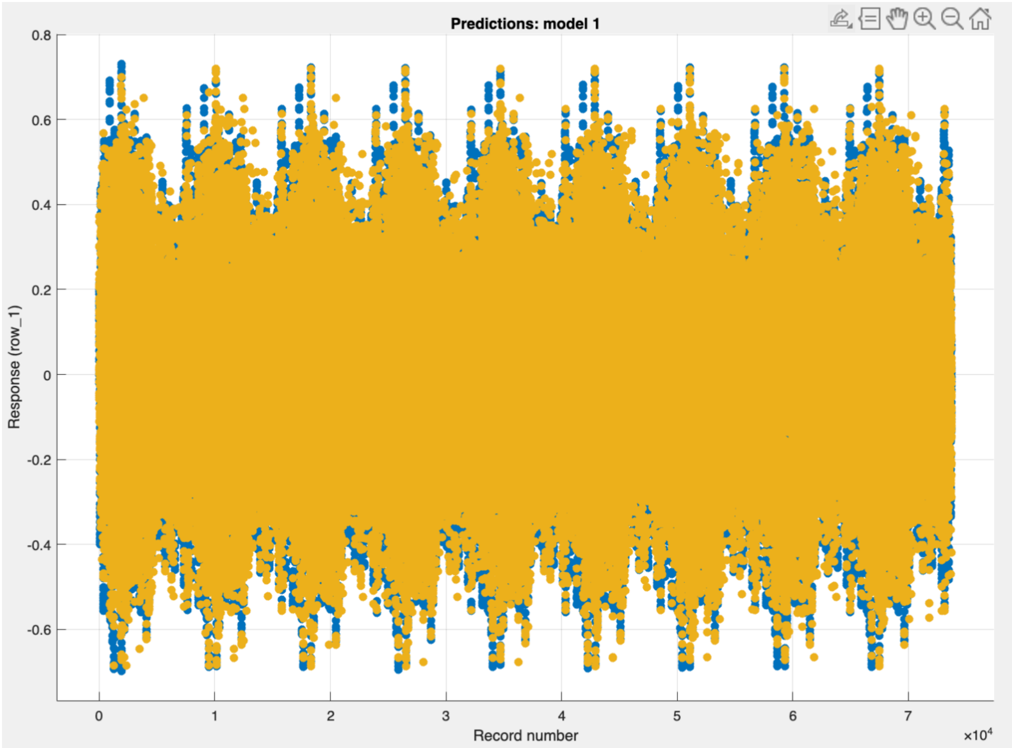


Figure 15: Response of Predicted vs Actual Acceleration 1

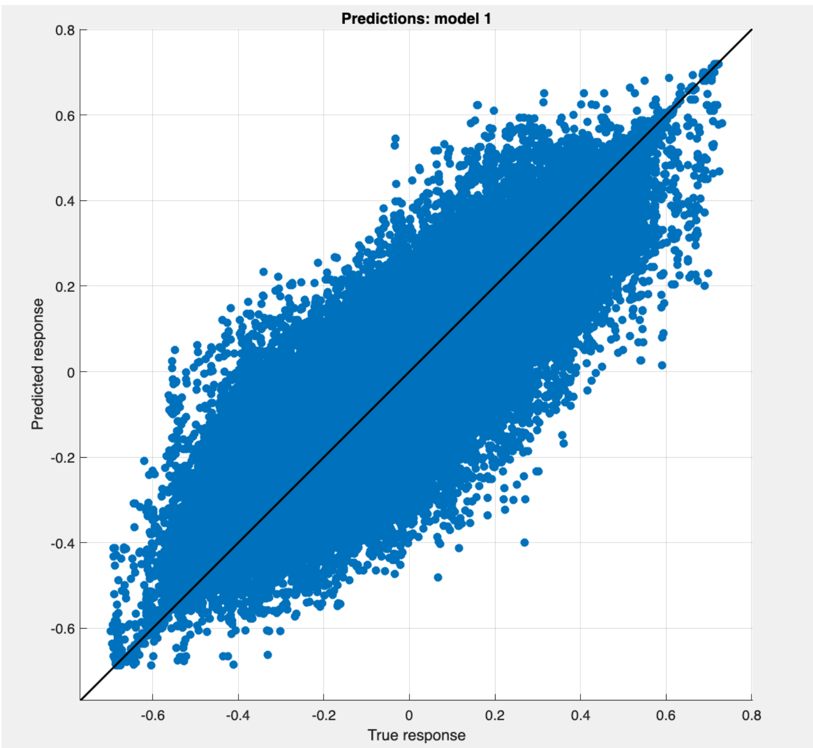


Figure 16: Validation Predicted vs Actual Acceleration 1

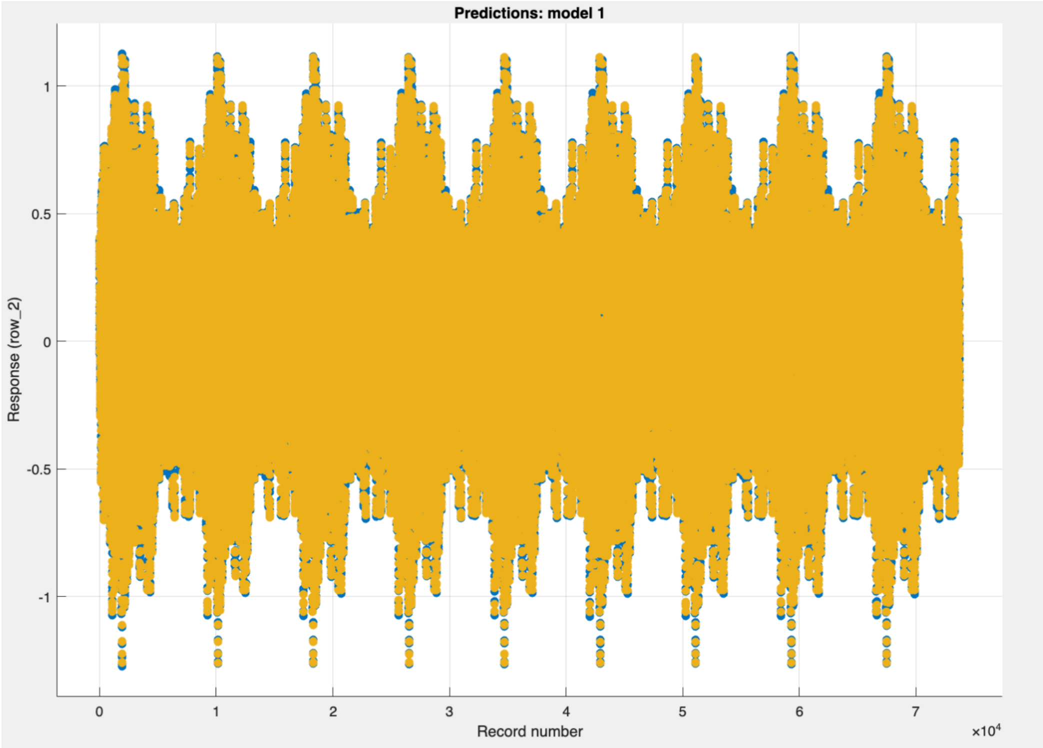


Figure 17: Response of Predicted vs Actual Acceleration 2

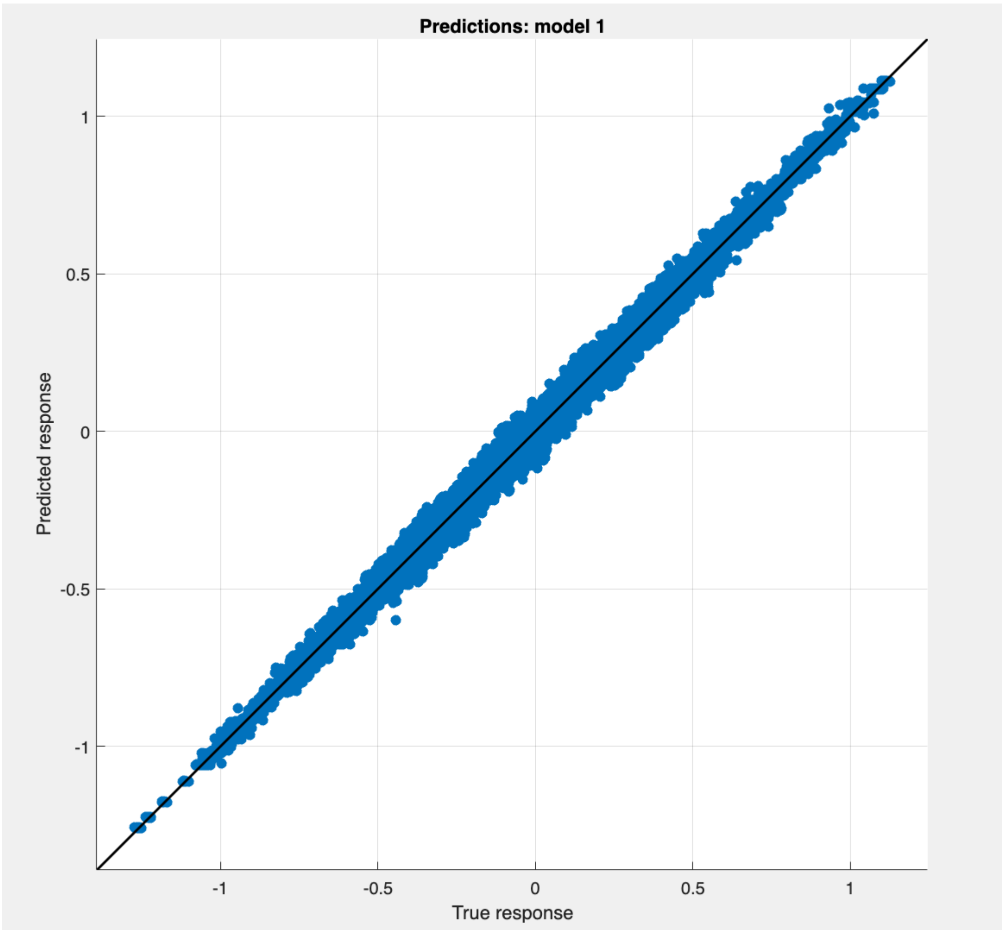


Figure 18: Validation Predicted vs Actual Acceleration 2

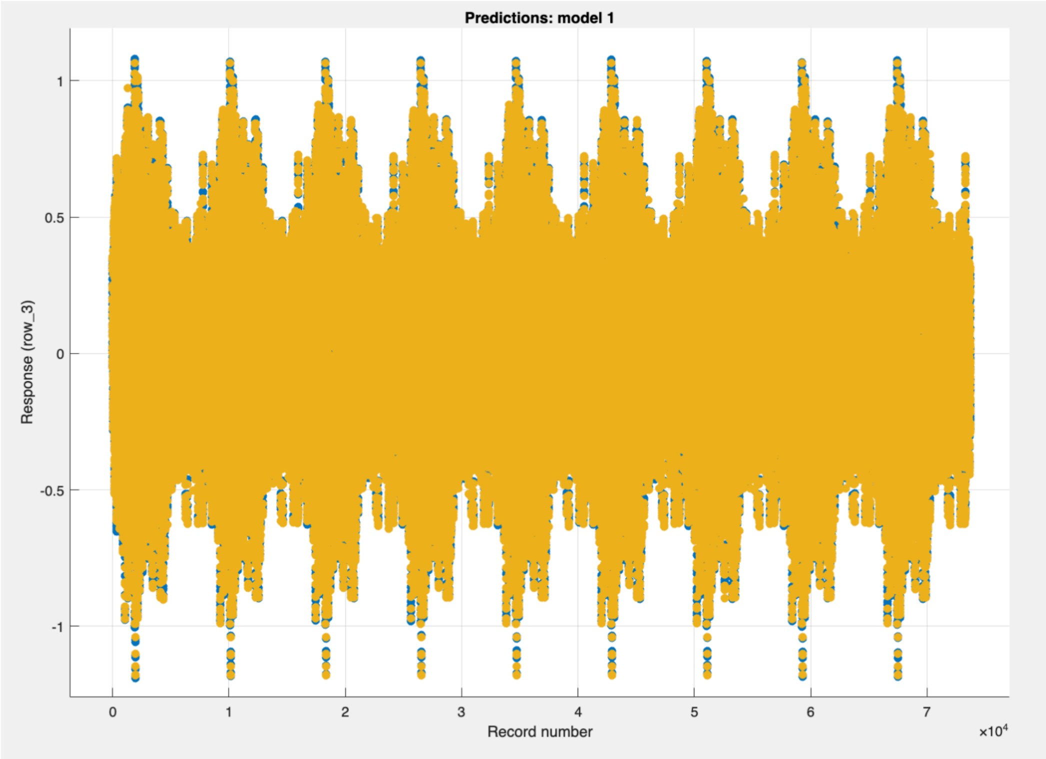


Figure 19: Response of Predicted vs Actual Acceleration 3

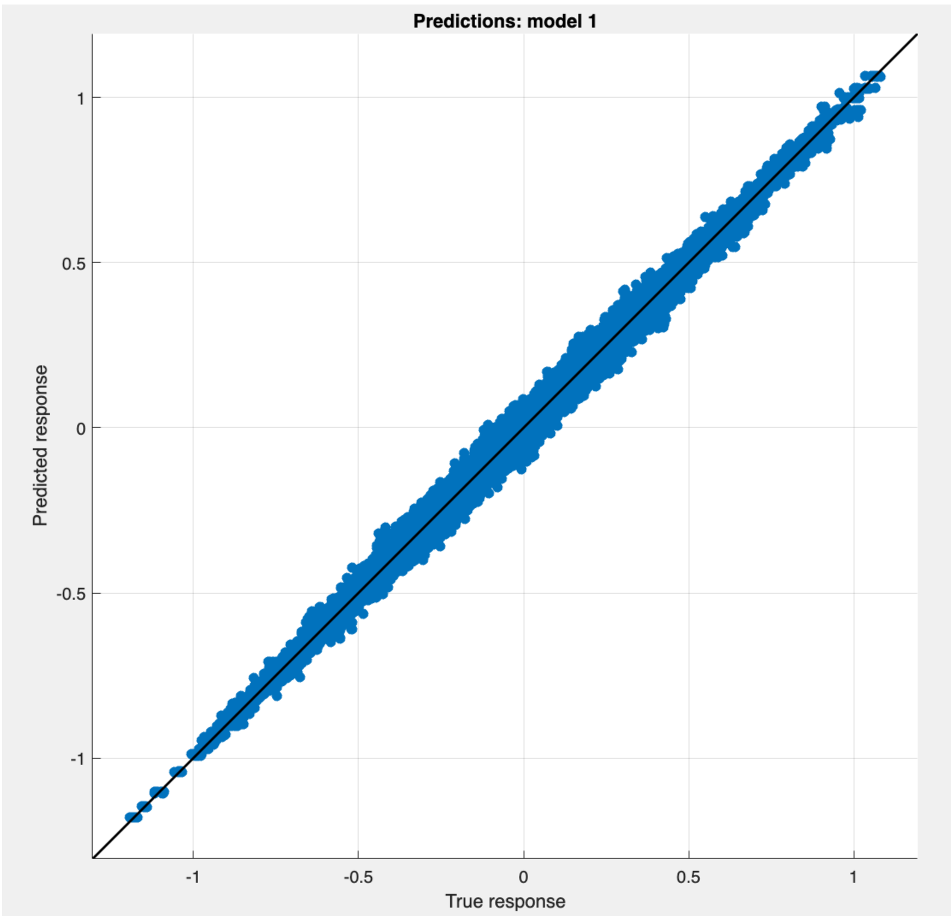


Figure 20: Validation Predicted vs Actual Acceleration 3

Table 11: Random Forest Training Results

Random Forest Training Results	
R-Value	0,916
MSE	3,97E-05
MAE	1,20E-05
Prediction Speed	1500000 obs/sec
Training Time	48.089 sec
Model Size	5 MB

These Random Forest training results are truly exceptional and indicative of a well-performing model. The high R-Value of 0.916 suggests a strong correlation between the features and the target variable, which is a crucial metric for regression models. Additionally, the low errors for Mean Squared Error (MSE) and Mean Absolute Error (MAE) indicate that the model's predictions are accurate and reliable.

The prediction speed of 1500000 observations per second is incredibly fast, showcasing the efficiency of the Random Forest algorithm in handling large datasets. Furthermore, the model size of 5 MB is relatively small, which makes it easy to deploy and integrate into production environments without consuming excessive resources.

Overall, these training results reflect a successful outcome and demonstrate the effectiveness of Random Forest in generating robust predictions. This model has the potential to provide valuable insights and make accurate forecasts, making it a valuable asset for any predictive modelling task.

5.2 Comparing Results

When we examine the results of the models used in the research conducted with the model used in this thesis in the table, it is clearly seen that the correlation coefficient we found showed a much better result than other models, but we achieved almost the same result with the Forer model. The result clearly shows that this model we used can be a very reliable model that can be used in aviation maintenance and is an example of a model that can be improved.

Table 12: Comparing Results

Model	References	R Results
SINDy	(NINNI, 2023)	• 0,9763
FOURIER NEURAL OPERATOR	(Park et al., 2023)	• 0,9780
NODE	(Gong et al., 2023)	• 0,9274
BLA	[(Csurcsia, Decuyper, Runacres, et al., 2023),(Csurcsia, Decuyper, Renczes, et al., 2023)]	• 0,9660
SAMI	(Csurcsia et al., 2020)	• 0,8823
NONLINEAR ARX	(Wigren & Schoukens, 2013)	• 0,9201
GAUSSIAN PROCESS MODELLING	(Perne et al., 2021)	• 0,8912
ANN		• 0,9268
LINEAR REGRESSION		• 0,8982
RANDOM FOREST		• 0,9160

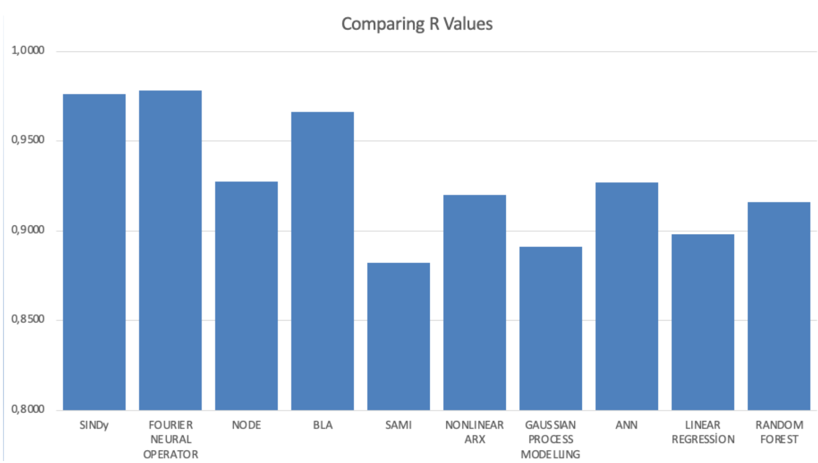


Figure 21: Comparing R Values

The issue at hand is that there are limited resources available for the growing population, leading to increased competition and potential conflicts over access to essential goods and services. This can manifest in various ways, such as disputes over land use, water rights, or access to healthcare and education. Additionally, unequal distribution of resources can exacerbate existing societal inequalities, creating further disparities between different groups within a community.

Furthermore, as the population continues to increase, the strain on the environment also escalates, leading to concerns about sustainability and long-term viability. Depletion of natural resources, pollution, and climate change are all pressing issues that must be addressed in order to ensure a healthier and more secure future for all.

In order to effectively address these challenges, it is essential for governments, organizations, and individuals to work together to find sustainable solutions that prioritize the well-being of both current and future generations. This may involve implementing policies to promote resource conservation, investing in renewable energy sources, fostering cooperation and dialogue between stakeholders, and promoting a mindset of shared responsibility and cooperation.

Chapter 6

CONCLUSION

6.1 Future Outlook

The future of aircraft maintenance lies in the continuous advancement and integration of Artificial Neural Networks and machine learning algorithms with other emerging technologies. The ongoing research and development in the field will lead to improved algorithms, enhanced interpretability, and better adaptation to evolving maintenance requirements. The integration of digital twins, augmented reality, and Internet of Things technologies will further revolutionize maintenance practices and enable real-time monitoring and decision-making.

Aircraft maintenance is a critical aspect of ensuring the safety and reliability of aviation operations. With advancements in technology, machine learning algorithms, specifically Artificial Neural Networks, have shown promise in improving maintenance processes. ANNs are a subset of machine learning algorithms that mimic the structure and functioning of the human brain, allowing them to learn from data and make predictions or decisions.

When applied to aircraft maintenance, ANNs can help analyse vast amounts of data from various sources, such as sensors, flight logs, maintenance records, and historical data. By processing this information, ANNs can detect patterns, identify anomalies, and predict potential failures or maintenance needs. This predictive capability enables airlines and maintenance organizations to move from reactive maintenance practices to proactive and predictive maintenance strategies.

ANNs can be utilized in different areas of aircraft maintenance. For example, they can be employed in engine health monitoring, where sensor data from aircraft engines is continuously monitored to detect deviations from normal behaviour and predict impending faults or maintenance requirements. Similarly, ANNs can be used in structural health monitoring, analysing data from strain gauges, accelerometers, and other sensors to assess the structural integrity of aircraft components and predict maintenance needs.

Moreover, ANNs can contribute to optimizing maintenance schedules and resource allocation. By analysing historical maintenance data and considering various factors such as flight cycles, component usage, and operational conditions, ANNs can provide recommendations on the optimal timing and type of maintenance activities. This approach helps minimize unscheduled maintenance events, reduce downtime, and optimize the allocation of maintenance resources.

To successfully implement ANNs in aircraft maintenance, it is essential to have high-quality data for training and validation. This data can include historical maintenance records, sensor data, fault reports, and other relevant information. Additionally, the development of accurate and reliable ANN models requires expertise in data pre-processing, feature selection, model architecture design, and algorithm tuning.
Several studies and research papers have been conducted on the topic of aircraft maintenance using ANNs and other machine learning algorithms. These studies

provide valuable insights into the potential benefits, challenges, and best practices of implementing such techniques in real-world maintenance scenarios. By referring to these studies, you can find specific examples, case studies, and experimental results that support the effectiveness of ANNs in aircraft maintenance.

6.2 Conclusion

This thesis gives a comprehensive review and analytic comparison of most recent design and optimization machine learning methods for aircraft maintenance as well as how ANN can be used optimally for an aircraft maintenance. In order to optimize aircraft maintenance system, various evaluation factors such as time and cost of maintenance are described and summarized. The consideration of some parameters such as safety, time of repair, time of replacement and maintenance cost is essentially in obtaining an optimal combination for the fastest and cheapest maintenance. These parameters will allow precise forecasting repair time, avoiding failures, ensuring the safety of aircraft and optimization of regular repair. Moreover, the availability of the required part and planning the maintenance have an impact on the aircraft maintenance optimization problem. Based on this review, maintenance schedule problem [S. Yan et al. (2004), Huang et al. (2010), Deng & Santos (2022), Andrade et al. (2021)], aircraft spare parts prediction problem [(Ghobbar & Friend, 2003; Regattieri et al., 2005)-Romeijnders et al. (2012), W. Wang (2012), Kazemi Zanjani & Nourelfath (2014), Capodieci et al. (2020)], predicting RUL of aircraft components [Azevedo et al. (2019b), T. Wang et al. (2008), Mathew et al. (2017),El Afia & Sarhani (2017),Azevedo et al. (2019a)], and predicting aircraft equipment failure [Celikmih et al. (2020), Dangut et al. (2021), Singh et al. (2020), Doğru et al. (2020)] are some of the problems that have been optimized by researchers using machine learning. It is important to accurately predict the RUL of components so as to maximize its lifetime and also know when it needs repair or replacement to avoid unseen failures. It is observed that artificial neural network is one of the main techniques used in developing predictive models to predict RUL, probably because of its ability to learn patterns and features in data to make accurate predictions (Nuhu et al., 2023), (Azevedo et al., 2019a). The unique feature of TF-IDF technique in extracting terms from document can be utilized to extract important features from data especially when solving classification problems (W. Yan & Zhou, 2017). Random Forest can handle both classification and regression issues with ease (Singh et al., 2020). Additionally, it can handle categorical as well as continuous data. Not to mention, it automates the procedure for adding missing values to data. Furthermore, Neural Network can save data throughout the whole network and be able to function with limited information. It has an excellent hazard and risk tolerance and also has a memory that is distributed. Mostly researchers used MATLAB and PYTHON to develop and test machine learning models.

Artificial Intelligence algorithms have recently been extensively applied for the optimization of aircraft maintenance system. Several researchers have proposed, analysed, and investigated the applications of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) based data analytics for predictive maintenance of aircraft systems, subsystems, and components. In order to inform more experts about the current research breakthroughs and to offer some direction for pertinent research in the field of aircraft maintenance this paper provided a comprehensive review of the ML techniques like Multilayer Perceptron (MLP), Logic Regression (LR), Random Forest (RF), Artificial Neural Network (ANN), Support Vector Regression (SVR), Linear Regression (LR), and other common ML techniques for their present implementation and potential applications in aircraft maintenance.

However, several gaps and research prospects need yet to be pursued. Alternative ML algorithms and architectures must be tested and explored for various aircraft maintenance applications. It is important to consider the most appropriate ML methodology that is best suited for a particular dataset and can give the best result for specific aircraft maintenance problems. It is therefore recommended that more real datasets may be made available in public domain so that more and more researchers can test their proposed ML algorithms on these real datasets and improve the efficiency and efficacy of their ML algorithms so that they can be extensively utilized for aircraft maintenance in the near future.

Furthermore, performance of any ML algorithm is heavily dependent on the training dataset. Major problems faced while using ML algorithms is having low quality or insufficient data, underfitting or overfitting of the training data etc. Moreover, implementation can take too much time and when data grows, there can be imperfections in the algorithm, that is why machine learning is still a tedious task and it is a long way till we can extensively and solely rely on ML data only for the design of maintenance operations. However, with the advent of faster computers, BDA platforms, Industrial internet of things (IIOT), Cloud Computation platforms, and significant advancements in data acquisitions systems, and prognostic health management tools, particularly Micro Electro-Mechanical System (MEMS) technologies, high-end microsensors for precision applications, it is anticipated that Machine Learning (ML) will play a far more substantial role in aircraft maintenance in years to come.

By leveraging the strengths of machine learning algorithms, the aviation industry can improve maintenance planning, reduce operational costs, and enhance overall safety.

Further research and experimentation are needed to explore and compare the performance of these algorithms on different datasets and specific maintenance tasks. By harnessing the power of machine learning, aircraft maintenance can be transformed into a more proactive, efficient, and data-driven process.

REFERENCES

A. U. Jan. (2020). *Opinion Mining and Summarization: A Comprehensive Review*. Inform. Comm. Tec. Robotic Applications. https://www.researchgate.net/publication/344313954_Opinion_Mining_and_Summarization_A_Comprehensive_Review

Ahmad, W. M. T. W., Ghani, N. L. A., & Drus, S. M. (2019). Data mining techniques for disease risk prediction model: A systematic literature review. *Advances in Intelligent Systems and Computing*, *843*, 40–46. https://doi.org/10.1007/978-3-319-99007-1_4/COVER

Ai, L., Soltangharaei, V., Bayat, M., Van Tooren, M., & Ziehl, P. (2021). Detection of impact on aircraft composite structure using machine learning techniques. *Measurement Science and Technology*, *32*(8), 084013. https://doi.org/10.1088/1361-6501/ABE790

Airbus. (2024). *accidentstats.airbus.com*. Airbus. https://accidentstats.airbus.com/

Akamine, M., & Ajmera, J. (2012). Decision tree-based acoustic models for speech recognition. *Eurasip Journal on Audio, Speech, and Music Processing*, *2012*(1), 1–8. https://doi.org/10.1186/1687-4722-2012-10/TABLES/3

Alessandrini, M., Falaschetti, L., Biagetti, G., Crippa, P., & Turchetti, C. (2022). Nonlinear Dynamic System Identification in the Spectral Domain Using Particle-

Bernstein Polynomials. *Electronics 2022, Vol. 11, Page 3100*, *11*(19), 3100. https://doi.org/10.3390/ELECTRONICS11193100

Altman, N., & Krzywinski, M. (2017). Points of Significance: Ensemble methods: Bagging and random forests. *Nature Methods*, *14*(10), 933–934. https://doi.org/10.1038/NMETH.4438

Andrade, P., Silva, C., Ribeiro, B., & Santos, B. F. (2021). Aircraft Maintenance Check Scheduling Using Reinforcement Learning. *Aerospace 2021, Vol. 8, Page 113*, *8*(4), 113. https://doi.org/10.3390/AEROSPACE8040113

Antolick, N. (2015). Machine Learning Algorithms for HUMS Improvement on Rotorcraft Components. *AHS Forum 71*. https://doi.org/10.4050/

Asuero, A. G., Sayago, A., & González, A. G. (2006). The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry*, *36*(1), 41–59. https://doi.org/10.1080/10408340500526766

Atasoy, B., Güllü, R., & Tan, T. (2012). Optimal inventory policies with non-stationary supply disruptions and advance supply information. *Decision Support Systems*, *53*(2), 269–281. https://doi.org/10.1016/J.DSS.2012.01.005

Awad, M., & Khanna, R. (2015). Efficient learning machines: Theories, concepts, and applications for engineers and system designers. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, 1–248. https://doi.org/10.1007/978-1-4302-5990-9/COVER

Azevedo, D., Ribeiro, B., & Cardoso, A. (2019a). Online simulation of methods to predict the remaining useful lifetime of aircraft components. *Proceedings of the 2019 5th Experiment at International Conference, Exp.at 2019*, 199–203. https://doi.org/10.1109/EXPAT.2019.8876574

Azevedo, D., Ribeiro, B., & Cardoso, A. (2019b). Web-based tool for predicting the Remaining Useful Lifetime of Aircraft Components. *Proceedings of the 2019 5th Experiment at International Conference, Exp.at 2019*, 231–232. https://doi.org/10.1109/EXPAT.2019.8876503

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/S10462-020-09896-5/METRICS

Blanco, V., Japón, A., & Puerto, J. (2018). Optimal arrangements of hyperplanes for multiclass classification. *Advances in Data Analysis and Classification 2019 14:1*, *14*(1), 175–199. https://doi.org/10.1007/s11634-019-00367-6

Bouarfa, S., Doğru, A., Arizar, R., Aydoğan, R., & Serafico, J. (2020). Towards automated aircraft maintenance inspection. A use case of detecting aircraft dents using mask r-cnn. *AIAA Scitech 2020 Forum*, *1 PartF*. https://doi.org/10.2514/6.2020-0389

Brandon Battles. (2003). *Aircraft Maintenance Costs: Significant but tricky | Aviation Pros*. AviationPros. https://www.aviationpros.com/aircraft/maintenance-providers/mro/article/10387195/aircraft-maintenance-costs-significant-but-tricky

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Brosnan Yuen. (2020). *(10) (PDF) Applications of Machine Learning*. https://www.researchgate.net/publication/344293508_Applications_of_Machine_Learning

Capodieci, A., Caricato, A., Carlucci, A. P., Ficarella, A., Mainetti, L., & Vergallo, C. (2020). Using different machine learning approaches to evaluate performance on spare parts request for aircraft engines. *E3S Web of Conferences*, *197*, 11014. https://doi.org/10.1051/E3SCONF/202019711014

Celikmih, K., Inan, O., & Uguz, H. (2020). Failure Prediction of Aircraft Equipment Using Machine Learning with a Hybrid Data Preparation Method. *Scientific Programming*, *2020*. https://doi.org/10.1155/2020/8616039

Charlotte Daniels. (2020). *Predictive Maintenance Analytics: Smarter, Safer & More Efficient Operations By Charlotte Daniels | Aerospace Tech Review*. Aerospace Tech Review. https://aerospacetechreview.com/predictive-maintenance-analytics-smarter-safer-more-efficient-operations-by-charlotte-daniels/

Çinar, Z. M., Nuhu, A. A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine Learning in Predictive Maintenance towards Sustainable Smart

Manufacturing in Industry 4.0. *Sustainability 2020, Vol. 12, Page 8211*, *12*(19), 8211. https://doi.org/10.3390/SU12198211

Clarke, L., Johnson, E., Nemhauser, G., & Zhu, Z. (1997). The aircraft rotation problem. *Annals of Operations Research*, *69*(0), 33–46. https://doi.org/10.1023/A:1018945415148/METRICS

Cline, B., Niculescu, R. S., Huffman, D., & Deckel, B. (2017). Predictive maintenance applications for machine learning. *Proceedings - Annual Reliability and Maintainability Symposium*. https://doi.org/10.1109/RAM.2017.7889679

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Csurcsia, P. Z., Decuyper, J., Renczes, B., & De Troyer, T. (2023). Nonlinear Modelling of an F16 Benchmark Measurement. *Conference Proceedings of the Society for Experimental Mechanics Series*, 49–60. https://doi.org/10.1007/978-3-031-04086-3_8/COVER

Csurcsia, P. Z., Decuyper, J., Runacres, M. C., Renczes, B., & De Troyer, T. (2023). Reducing Black-Box Nonlinear State-Space Models: A Real-Life Case Study. *SSRN*. https://doi.org/10.2139/SSRN.4606907

Csurcsia, P. Z., Peeters, B., Schoukens, J., & De Troyer, T. (2020). Simplified Analysis for Multiple Input Systems: A Toolbox Study Illustrated on F-16 Measurements. *Vibration*, *3*(2), 70–84. https://doi.org/10.3390/vibration3020007

D. K. Bhattacharya. (2017). *Machine Learning: A Turning Point For Predictive Maintenance?* OliverWyman Velocity. https://www.oliverwyman.com/our-expertise/insights/2017/sep/oliver-wyman-transport-and-logistics-2017/operations/machine-learning-a-turning-point-for-predictive-maintenance.html

Dangut, M. D., Skaf, Z., & Jennions, I. K. (2021). An integrated machine learning model for aircraft components rare failure prognostics with log-based dataset. *ISA Transactions*, *113*, 127–139. https://doi.org/10.1016/J.ISATRA.2020.05.001

Deng, Q., & Santos, B. F. (2022). Lookahead approximate dynamic programming for stochastic aircraft maintenance check scheduling optimization. *European Journal of Operational Research*, *299*(3), 814–833. https://doi.org/10.1016/J.EJOR.2021.09.019

Dijkstra, M. C., Kroon, L. G., Salomon, M., Nunen, J. A. E. E. van, & Wassenhove, L. N. Van. (1994). Planning the Size and Organization of KLM's Aircraft Maintenance Personnel. *Https://Doi.Org/10.1287/Inte.24.6.47*, *24*(6), 47–58. https://doi.org/10.1287/INTE.24.6.47

Dijkstra, M. C., Kroon, L. G., van Nunen, J. A. E. E., & Salomon, M. (1991). A DSS for capacity planning of aircraft maintenance personnel. *International Journal of Production Economics*, *23*(1–3), 69–78. https://doi.org/10.1016/0925-5273(91)90049-Y

Dinis, D., & Barbosa-Póvoa, A. P. (2015). On the Optimization of Aircraft Maintenance Management. *Studies in Big Data*, *15*, 49–57. https://doi.org/10.1007/978-3-319-24154-8_7/COVER

Doğru, A., Bouarfa, S., Arizar, R., & Aydoğan, R. (2020). Using Convolutional Neural Networks to Automate Aircraft Maintenance Visual Inspection. *Aerospace 2020, Vol. 7, Page 171*, *7*(12), 171. https://doi.org/10.3390/AEROSPACE7120171

Dos Santos, T., Ferreira, F. J. T. E., Pires, J. M., & Damasio, C. (2017). Stator winding short-circuit fault diagnosis in induction motors using random forest. *2017 IEEE International Electric Machines and Drives Conference, IEMDC 2017*. https://doi.org/10.1109/IEMDC.2017.8002350

El Afia, A., & Sarhani, M. (2017). Particle swarm optimization for model selection of aircraft maintenance predictive models. *ACM International Conference Proceeding Series*, *Part F129474*. https://doi.org/10.1145/3090354.3090402

Elettronici Course, S., & Tania Cerquitelli, P. (2017). *POLITECNICO DI TORINO Electronic Engineering Master Degree Creation of a Machine Learning model for the Predictive Maintenance of an engine equipped with a rotating shaft*.

Essam Al Daoud. (2019). Comparison between XGBoost, LightGBM and CatBoost Using a Home Credit Dataset. *World Academy of Science, Engineering and TechnologyInternational Journal of Computer and Information Engineering*. https://doi.org/10.5281/ZENODO.3607805

Fernandes, S., Antunes, M., Santiago, A. R., Barraca, J. P., Gomes, D., & Aguiar, R. L. (2020). Forecasting Appliances Failures: A Machine-Learning Approach to Predictive Maintenance. *Information 2020, Vol. 11, Page 208*, *11*(4), 208. https://doi.org/10.3390/INFO11040208

G. Kavita. (n.d.). *What is Term-Frequency? - Kavita Ganesan, PhD*. Ganesan. Retrieved April 8, 2024, from https://kavita-ganesan.com/what-is-term-frequency/#.Y7zno3ZBzIU

Ghazal, T. M., Hussain, M. Z., Said, R. A., Nadeem, A., Hasan, M. K., Ahmad, M., Khan, M. A., & Naseem, M. T. (2021). Performances of K-Means Clustering Algorithm with Different Distance Metrics. *Intelligent Automation & Soft Computing*, *30*(2), 735–742. https://doi.org/10.32604/IASC.2021.019067

Ghobbar, A. A., & Friend, C. H. (2003). Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. *Computers & Operations Research*, *30*(14), 2097–2114. https://doi.org/10.1016/S0305-0548(02)00125-9

Godoy, W. F., da Silva, I. N., Goedtel, A., Palácios, R. H. C., & Lopes, T. D. (2016). Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter. *IET Electric Power Applications*, *10*(5), 430–439. https://doi.org/10.1049/IET-EPA.2015.0469

Gong, S., Meng, Q., Wang, Y., Wu, L., Chen, W., Ma, Z., & Liu, T. Y. (2023). Incorporating NODE with pre-trained neural differential operator for learning dynamics. *Neurocomputing*, *528*, 48–58. https://doi.org/10.1016/J.NEUCOM.2023.01.040

Grafmüller, M., Beyerer, J., & Kroschel, K. (2010). Decision tree classifier for character recognition combining support vector machines and artificial neural networks. *Https://Doi.Org/10.1117/12.860500*, *7799*, 99–106. https://doi.org/10.1117/12.860500

Gu, J., Zhang, G., & Li, K. W. (2015). Efficient aircraft spare parts inventory management under demand uncertainty. *Journal of Air Transport Management*, *42*, 101–109. https://doi.org/10.1016/J.JAIRTRAMAN.2014.09.006

Gutierrez, R. S., Solis, A. O., & Mukhopadhyay, S. (2008). Lumpy demand forecasting using neural networks. *International Journal of Production Economics*, *111*(2), 409–420. https://doi.org/10.1016/J.IJPE.2007.01.007

Hobbs, A. (2021). Aircraft Maintenance and Inspection. *International Encyclopedia of Transportation: Volume 1-7*, *2*, 25–33. https://doi.org/10.1016/B978-0-08-102671-7.10103-4

Huang, Z., Chang, W., Xiao, Y., & Liu, R. (2010). Optimizing human resources allocation on aircraft maintenance with predefined sequence. *2010 International Conference on Logistics Systems and Intelligent Management, ICLSIM 2010*, *2*, 1018–1022. https://doi.org/10.1109/ICLSIM.2010.5461109

I. Dave. (2018). *Similarity Machine Learning | Nearest Neighbor ML | simMachines*. Simmachines. https://simmachines.com/similarity-based-machine-learning-provides-ai-transparency-trust/

Jalawkhan, M. S., & Mustafa, T. K. (2021). Anomaly Detection in Flight Data Using the Naïve Bayes Classifier. *7th International Conference on Contemporary Information Technology and Mathematics, ICCITM 2021*, 26–30. https://doi.org/10.1109/ICCITM53167.2021.9677655

Jason Browniee. (2023). *Linear Regression for Machine Learning - MachineLearningMastery.com*. Machine Learning Mastery. https://machinelearningmastery.com/linear-regression-for-machine-learning/

Karaoğlu, U., Mbah, O., & Zeeshan, Q. (2022). Applications of Machine Learning in Aircraft Maintenance. *Journal of Engineering Management and Systems Engineering*, *2*(1), 76–95. https://doi.org/10.56578/jemse020105

Kazemi Zanjani, M., & Nourelfath, M. (2014). Integrated spare parts logistics and operations planning for maintenance service providers. *International Journal of Production Economics*, *158*, 44–53. https://doi.org/10.1016/J.IJPE.2014.07.012

Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. *International Journal of Production Economics*, *143*(1), 198–206. https://doi.org/10.1016/J.IJPE.2013.01.009

Luna, J. M., Gennatas, E. D., Ungar, L. H., Eaton, E., Diffenderfer, E. S., Jensen, S. T., Simone, C. B., Friedman, J. H., Solberg, T. D., & Valdes, G. (2019). Building more accurate decision trees with the additive tree. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(40), 19887–19893. https://doi.org/10.1073/PNAS.1816748116/SUPPL_FILE/PNAS.1816748116.SAPP.PDF

Maarten Schoukens. (2017). *F-16 aircraft benchmark based on ground vibration test data*. Workshop on Nonlinear System Identification Benchmarks. https://www.researchgate.net/publication/311776429_F-16_aircraft_benchmark_based_on_ground_vibration_test_data

Mathew, V., Toby, T., Singh, V., Rao, B. M., & Kumar, M. G. (2017). Prediction of Remaining Useful Lifetime (RUL) of turbofan engine using machine learning. *IEEE International Conference on Circuits and Systems, ICCS 2017*, *2018-January*, 306–311. https://doi.org/10.1109/ICCS1.2017.8326010

MathWorks. (2024). *Selecting a Multilayer Neural Network Training Function - MATLAB & Simulink - MathWorks United Kingdom*. MathWorks. https://ww2.mathworks.cn/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html

McCartney, M., Haeringer, M., & Polifke, W. (2019). Comparison of machine learning algorithms in the interpolation and extrapolation of flame describing functions. *Proceedings of the ASME Turbo Expo*, *4B-2019*. https://doi.org/10.1115/GT2019-91319

Murty, M. N., & Devi, V. S. (2011). *Nearest Neighbour Based Classifiers*. 48–85. https://doi.org/10.1007/978-0-85729-495-1_3

Nguyen, K. T. P., & Medjaher, K. (2019). A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliability Engineering & System Safety*, *188*, 251–262. https://doi.org/10.1016/J.RESS.2019.03.018

NINNI, D. (2023). *Physics-Informed Machine Learning for High-Fidelity Synthetic Data Generation*. https://thesis.unipd.it/handle/20.500.12608/47364

Nuhu, A. A., Zeeshan, Q., Safaei, B., & Shahzad, M. A. (2023). Machine learning-based techniques for fault diagnosis in the semiconductor manufacturing process: a comparative study. *Journal of Supercomputing*, *79*(2), 2031–2081. https://doi.org/10.1007/S11227-022-04730-X/METRICS

Nurhadiyanto, Nugroho, S. M. S., & Setijadi, E. (2019). Classification of Aircraft Inspection Result Using K-Nearest Neighbors. *Proceedings - 2019 International Seminar on Intelligent Technology and Its Application, ISITIA 2019*, 234–238. https://doi.org/10.1109/ISITIA.2019.8937081

P Korvesis. (2017). Machine Learning for Predictive Maintenance in Aviation. *HAL-Pastel*.

Paluszek, M., & Thomas, S. (2019). MATLAB machine learning recipes: A problem-solution approach, Second edition. *MATLAB Machine Learning Recipes: A*

Problem-Solution Approach, Second Edition, 1–347. https://doi.org/10.1007/978-1-4842-3916-2/COVER

Park, Y., Choi, J., Yoon, C., Song, C. hoon, & Kang, M. (2023). *Learning PDE Solution Operator for Continuous Modeling of Time-Series*. http://arxiv.org/abs/2302.00854

Partha Adhikari. (2018). *Machine Learning based Data Driven Diagnostics & Prognostics Framework for Aircraft Predictive Maintenance*. Creative Commons. https://www.ndt.net/search/docs.php3?id=23875

Patrick KY. (2020). *EASA Artificial Intelligence Roadmap 1.0 published - A human-centric approach to AI in aviation | EASA*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/newsroom-and-events/news/easa-artificial-intelligence-roadmap-10-published

Perne, M., Stepancic, M., & Kocijan, J. (2021). Gaussian Process Modelling of the F-16 Ground Vibration Test Benchmark: Data Selection Case Study. *IFAC-PapersOnLine*, *54*(4), 56–61. https://doi.org/10.1016/J.IFACOL.2021.10.010

Raschka, S., Patterson, J., & Nolet, C. (2020). Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. *Information 2020, Vol. 11, Page 193*, *11*(4), 193. https://doi.org/10.3390/INFO11040193

Razzaghi, P., Tabrizian, A., Guo, W., Chen, S., Taye, A., Thompson, E., Bregeon, A., Baheri, A., & Wei, P. (2022). *A Survey on Reinforcement Learning in Aviation Applications*. https://arxiv.org/abs/2211.02147v2

Regattieri, A., Gamberi, M., Gamberini, R., & Manzini, R. (2005). Managing lumpy demand for aircraft spare parts. *Journal of Air Transport Management*, *11*(6), 426–431. https://doi.org/10.1016/J.JAIRTRAMAN.2005.06.003

Rengasamy, D., Morvan, H. P., & Figueredo, G. P. (2018). Deep Learning Approaches to Aircraft Maintenance, Repair and Overhaul: A Review. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-November*, 150–156. https://doi.org/10.1109/ITSC.2018.8569502

Romeijnders, W., Teunter, R., & Van Jaarsveld, W. (2012). A two-step method for forecasting spare parts demand using information on component repairs. *European Journal of Operational Research*, *220*(2), 386–393. https://doi.org/10.1016/J.EJOR.2012.01.019

Ruan, J. H., Wang, Z. X., Chan, F. T. S., Patnaik, S., & Tiwari, M. K. (2021). A reinforcement learning-based algorithm for the aircraft maintenance routing problem. *Expert Systems with Applications*, *169*, 114399. https://doi.org/10.1016/J.ESWA.2020.114399

S. Paul. (2010). *Application of Artificial Neural Networks in Aircraft Maintenance, Repair and Overhaul Solutions*. ArXiv.

https://www.researchgate.net/publication/45896609_Application_of_Artificial_Neural_Networks_in_Aircraft_Maintenance_Repair_and_Overhaul_Solutions

Safaei, N., Banjevic, D., & Jardine, A. K. S. (2011). Workforce-constrained maintenance scheduling for military aircraft fleet: A case study. *Annals of Operations Research*, *186*(1), 295–316. https://doi.org/10.1007/S10479-011-0885-4/METRICS

Sano, H. H., & Berton, L. (2022). Application of Deep Learning Models for Aircraft Maintenance. *Anais Do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, 787–797. https://doi.org/10.5753/ENIAC.2022.227575

Sarkon, G. K., Safaei, B., Kenevisi, M. S., Arman, S., & Zeeshan, Q. (2022). State-of-the-Art Review of Machine Learning Applications in Additive Manufacturing; from Design to Manufacturing and Property Control. *Archives of Computational Methods in Engineering 2022 29:7*, *29*(7), 5663–5721. https://doi.org/10.1007/S11831-022-09786-9

Seale, M., Hines, A., Nabholz, G., Ruvinsky, A., Eslinger, O., Rigoni, N., & Vega-Maisonet, L. (2019). Approaches for Using Machine Learning Algorithms with Large Label Sets for Rotorcraft Maintenance. *IEEE Aerospace Conference Proceedings*, *2019-March*. https://doi.org/10.1109/AERO.2019.8742027

Singh, D., Kumar, M., Arya, K. V., & Kumar, S. (2020). Aircraft Engine Reliability Analysis using Machine Learning Algorithms. *2020 IEEE 15th International*

Conference on Industrial and Information Systems, ICIIS 2020 - Proceedings, 443–448. https://doi.org/10.1109/ICIIS51140.2020.9342675

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, *24*(1), 12–18. https://doi.org/10.11613/BM.2014.003/FULLARTICLE

Wang, T., Yu, J., Siegel, D., & Lee, J. (2008). A similarity-based prognostics approach for remaining useful life estimation of engineered systems. *2008 International Conference on Prognostics and Health Management, PHM 2008*. https://doi.org/10.1109/PHM.2008.4711421

Wang, W. (2012). A stochastic model for joint spare parts inventory and planned maintenance optimisation. *European Journal of Operational Research*, *216*(1), 127–139. https://doi.org/10.1016/J.EJOR.2011.07.031

Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, *21*(6), 2560–2574. https://doi.org/10.1016/J.YMSSP.2006.12.007

Wigren, T., & Schoukens, J. (2013). Three free data sets for development and benchmarking in nonlinear system identification. *2013 European Control Conference, ECC 2013*, 2933–2938. https://doi.org/10.23919/ECC.2013.6669201

Yan, S., Yang, T. H., & Chen, H. H. (2004). Airline short-term maintenance manpower supply planning. *Transportation Research Part A: Policy and Practice*, *38*(9–10), 615–642. https://doi.org/10.1016/J.TRA.2004.03.005

Yan, W., & Zhou, J. H. (2017). Predictive modeling of aircraft systems failure using term frequency-inverse document frequency and random forest. *IEEE International Conference on Industrial Engineering and Engineering Management*, *2017-December*, 828–831. https://doi.org/10.1109/IEEM.2017.8290007

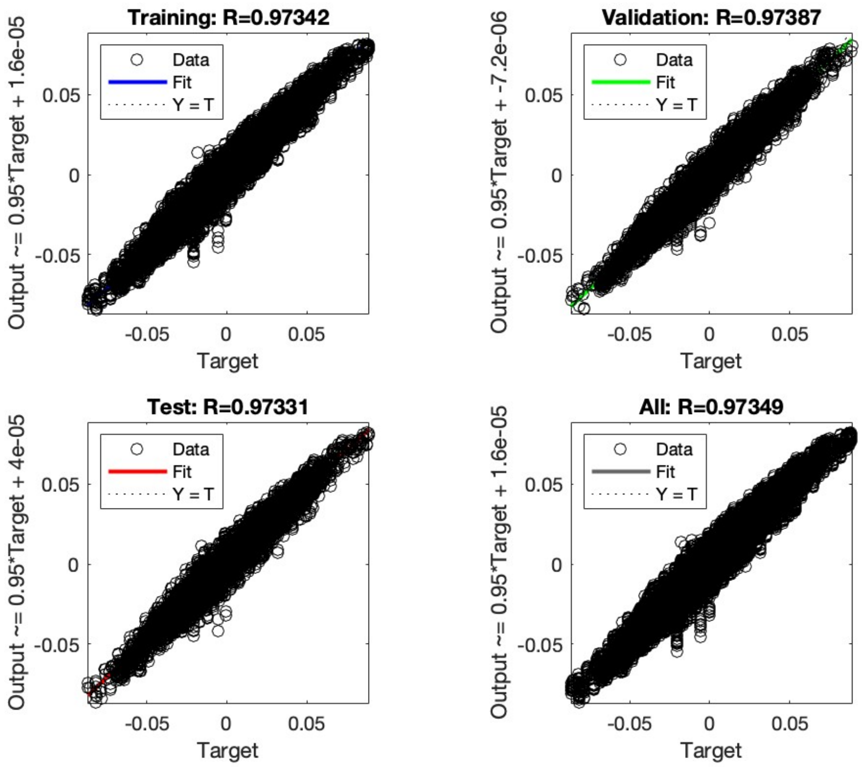
Zenisek, J., Holzinger, F., & Affenzeller, M. (2019). Machine learning based concept drift detection for predictive maintenance. *Computers & Industrial Engineering*, *137*, 106031. https://doi.org/10.1016/J.CIE.2019.106031

APPENDICES

Appendix A: F16 Data

	Acceleration 1	Acceleration 2	Acceleration 3	Volatge	Force	Fs
1	-0,0184	-0,0496	-0,0463	-0,0005	0,9571	400
2	-0,0184	-0,0332	-0,0275	0,0027	-0,5240	400
3	-0,0160	-0,0195	-0,0126	0,0061	-20148,0000	400
4	-0,0211	-0,0041	0,0126	0,0094	-0,4907	400
5	-0,0154	0,0115	0,0254	0,0126	-5,0049	400
6	-0,0131	0,0271	0,0402	0,0156	-6,2549	400
7	-0,0145	0,0462	0,0658	0,0182	-7,3773	400
8	-0,0022	0,0647	0,0824	0,0205	-8,3348	400
9	0,0035	0,0855	0,1061	0,0222	-8,9568	400
10	0,0136	0,1111	0,1301	0,0235	-9,4585	400
	.	.	.	.	.	.
	.	.	.	.	.	.
	.	.	.	.	.	.
	.	.	.	.	.	.
73724	0,0034	-0,0459	-0,0522	-0,0114	5,9443	400
73725	0,0086	-0,0290	-0,0338	-0,0101	5,5052	400
73726	0,0113	-0,0103	-0,0145	-0,0084	4,6959	400
73727	0,0160	0,0020	-0,0024	-0,0061	3,6719	400
73728	0,0094	0,0128	0,0147	-0,0034	2,4951	400

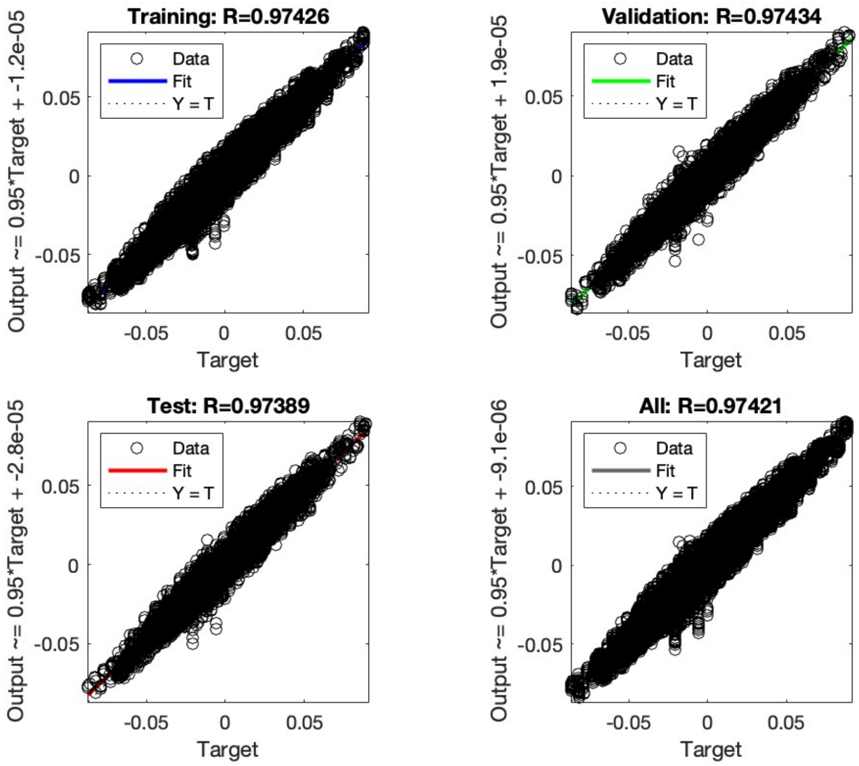
Appendix B: Analysis Results



Appendix 1 Analysis Result of 60-20-20-5

	Observations	MSE	R
Training	44236	3.2835e-05	0.9734
Validation	14746	3.2321e-05	0.9739
Test	14746	3.3062e-05	0.9733

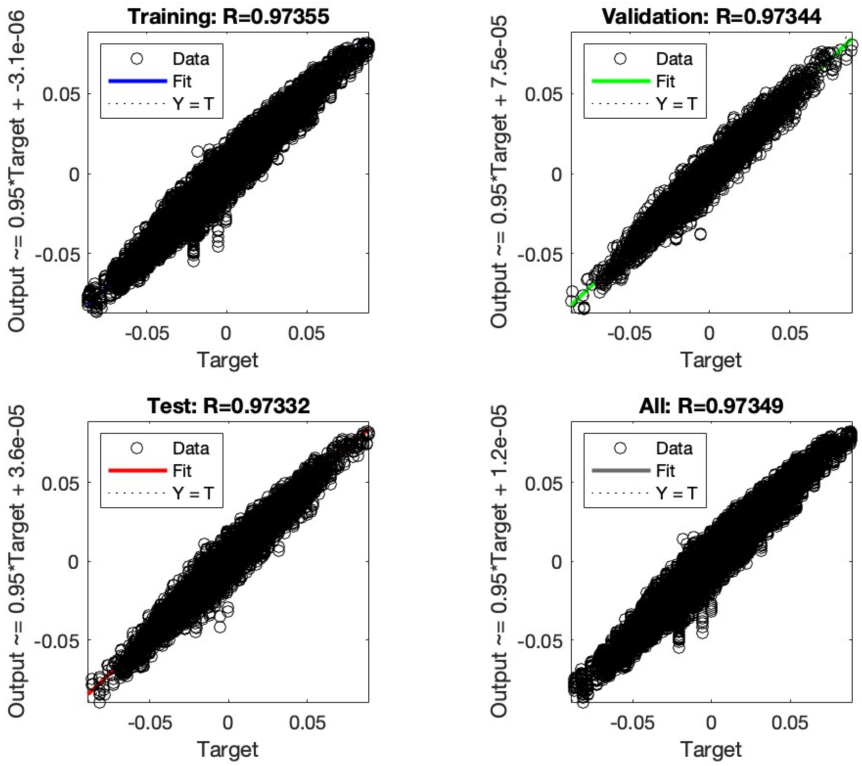
Appendix 2 MatLab Result of 60-20-20-5



Appendix 3 *Analysis Result of 60-20-20-10*

	Observations	MSE	R
Training	51609	3.2414e-05	0.9739
Validation	7373	3.2198e-05	0.9732
Test	14746	3.2262e-05	0.9741

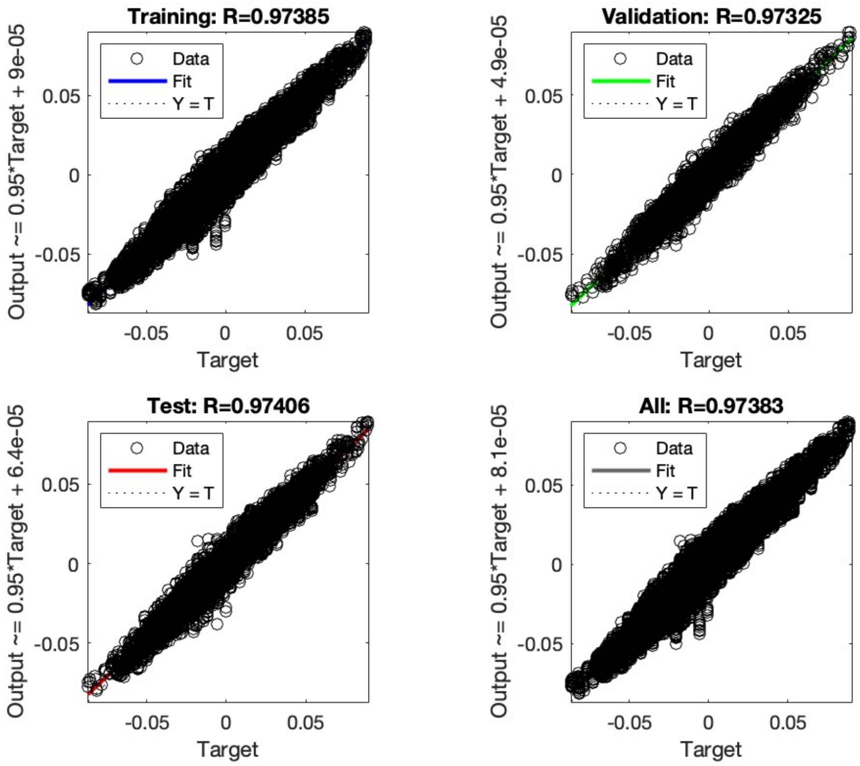
Appendix 4 *MatLab Result of 60-20-20-10*



Appendix 5 Analysis Result of 70-10-20-5

	Observations	MSE	R
Training	51609	3.2721e-05	0.9736
Validation	7373	3.2517e-05	0.9734
Test	14746	3.3055e-05	0.9733

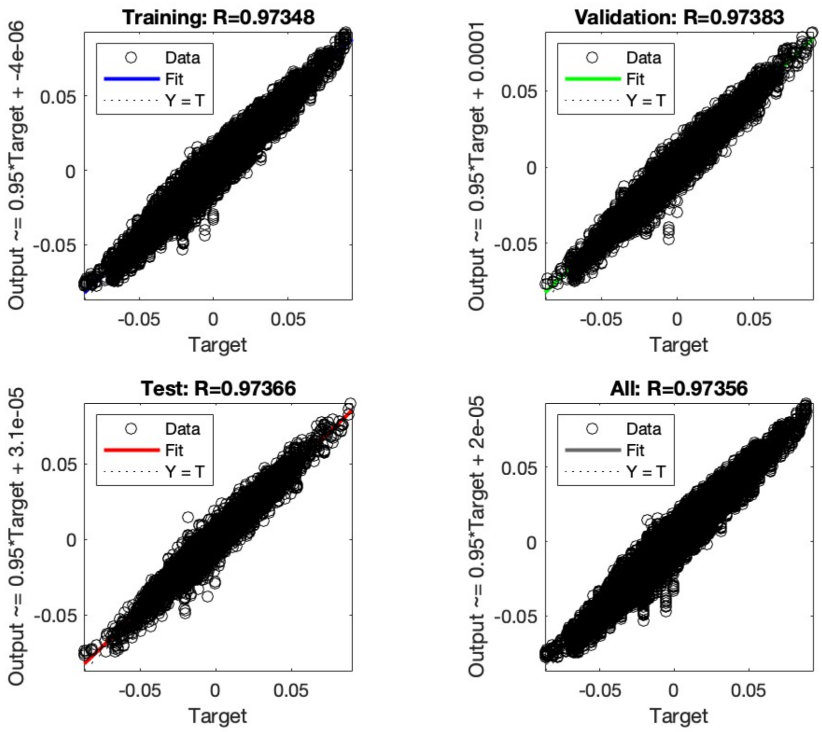
Appendix 6 MatLab Result of 70-10-20-5



Appendix 7 Analysis Result of 70-10-20-10

	Observations	MSE	R
Training	51609	3.2414e-05	0.9739
Validation	7373	3.2198e-05	0.9732
Test	14746	3.2262e-05	0.9741

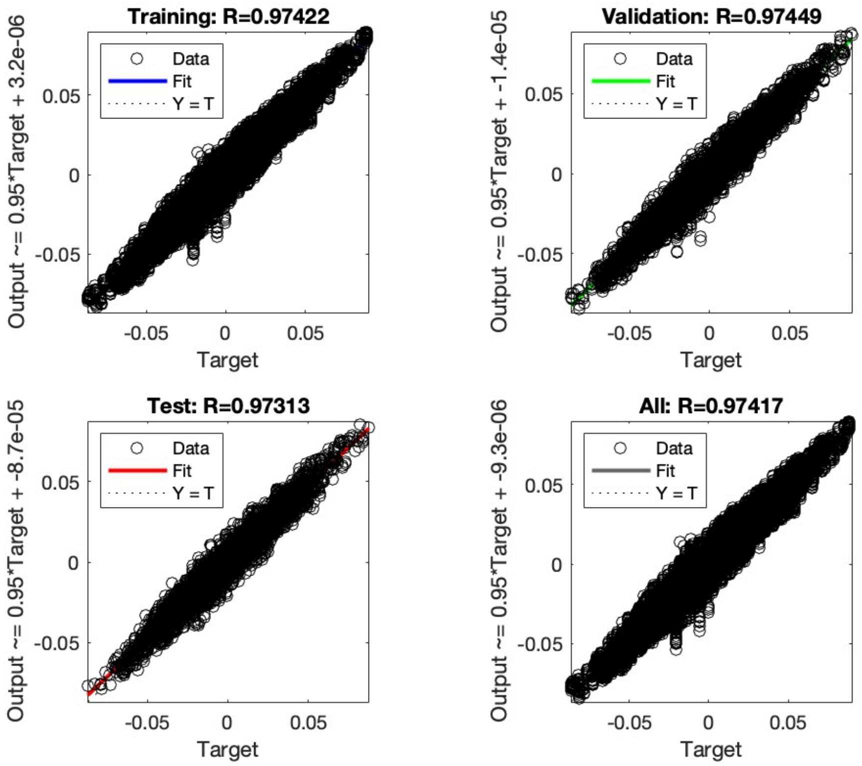
Appendix 8 MatLab Result of 70-10-20-10



Appendix 9 Analysis Result of 70-20-10-5

	Observations	MSE	R
Training	51609	3.2586e-05	0.9735
Validation	14746	3.2634e-05	0.9738
Test	7373	3.3453e-05	0.9737

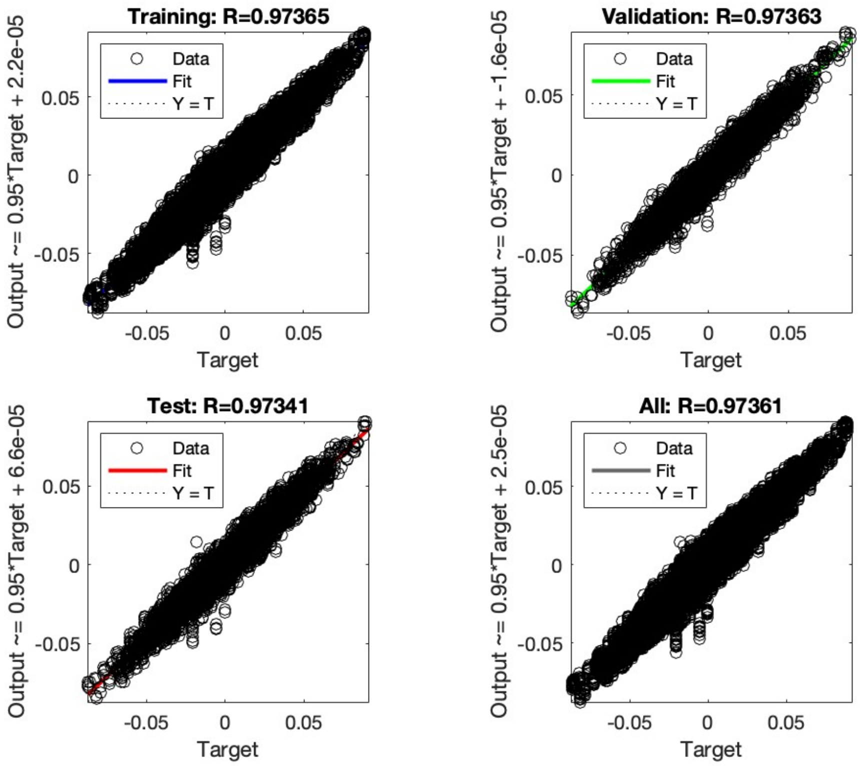
Appendix 10 Matlab Result 70-20-10-5



Appendix 11 Analysis Result 70-20-10-10

	Observations	MSE	R
Training	51609	3.1874e-05	0.9742
Validation	14746	3.1892e-05	0.9745
Test	7373	3.2565e-05	0.9731

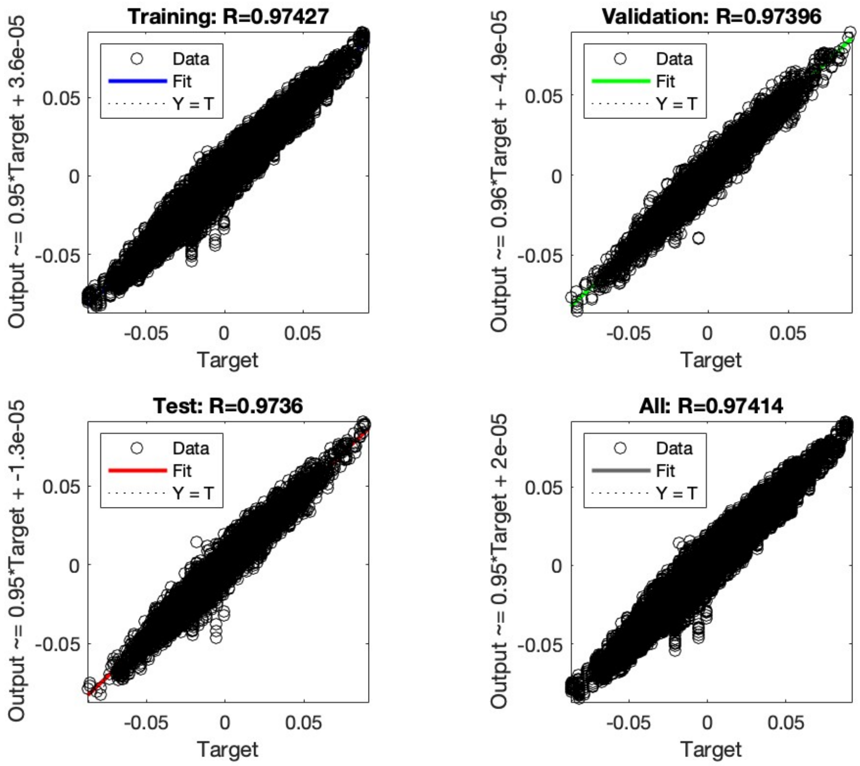
Appendix 12 Matlab Result 70-20-10-10



Appendix 13 Analysis Result 75-10-15-5

	Observations	MSE	R
Training	55296	3.2509e-05	0.9736
Validation	7373	3.2381e-05	0.9736
Test	11059	3.3385e-05	0.9734

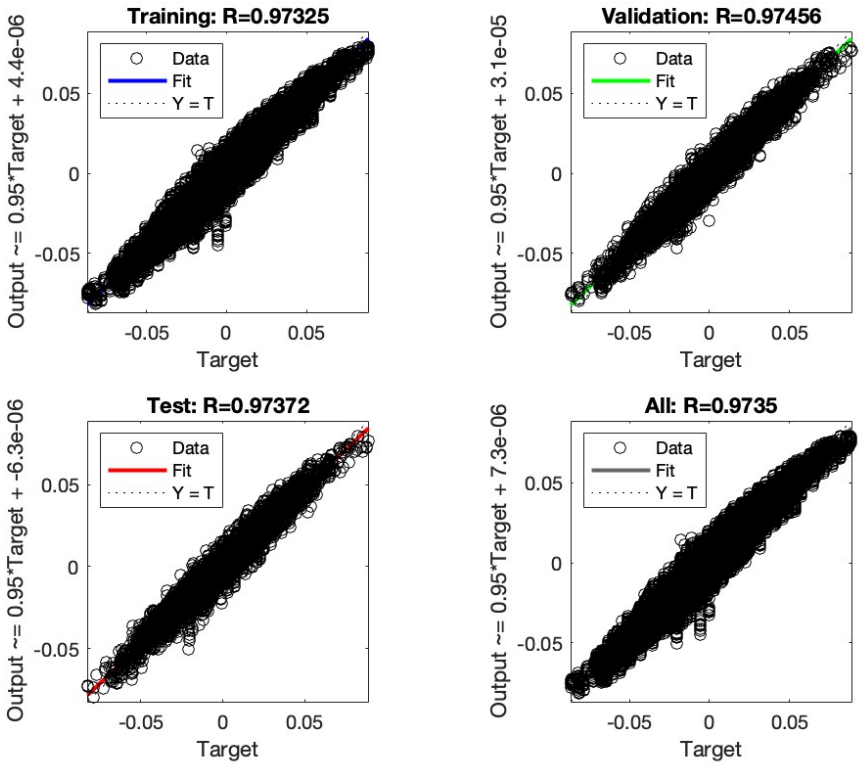
Appendix 14 MatLab Result 75-10-15-5



Appendix 15 Analysis Result 75-10-15-10

	Observations	MSE	R
Training	55296	3.1866e-05	0.9743
Validation	7373	3.2630e-05	0.9740
Test	11059	3.2138e-05	0.9736

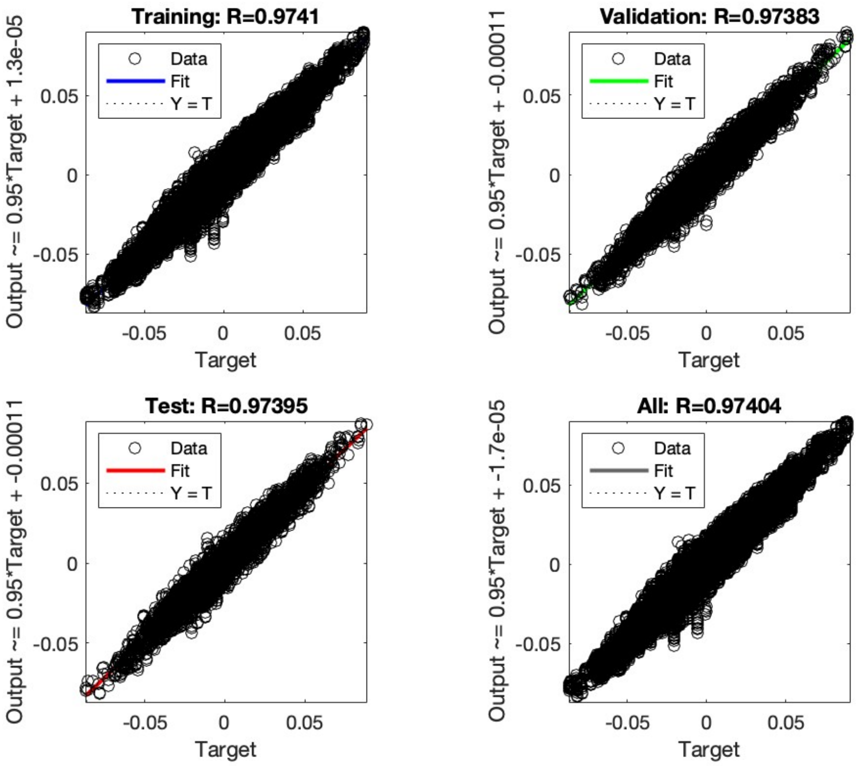
Appendix 16 MatLab Result 75-10-15-10



Appendix 17 Analysis Result 75-15-10-5

	Observations	MSE	R
Training	55296	3.2888e-05	0.9733
Validation	11059	3.2456e-05	0.9746
Test	7373	3.2263e-05	0.9737

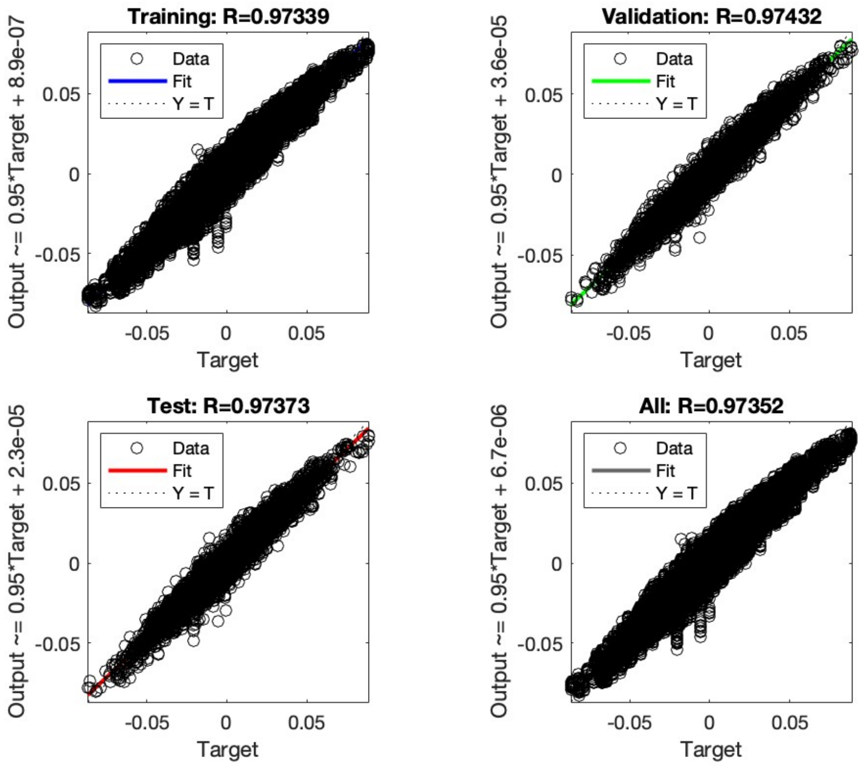
Appendix 18 MatLab results 75-15-10-5



Appendix 19 Analysis Result 75-15-10-10

	Observations	MSE	R
Training	55296	3.2143e-05	0.9741
Validation	11059	3.2107e-05	0.9738
Test	7373	3.1740e-05	0.9739

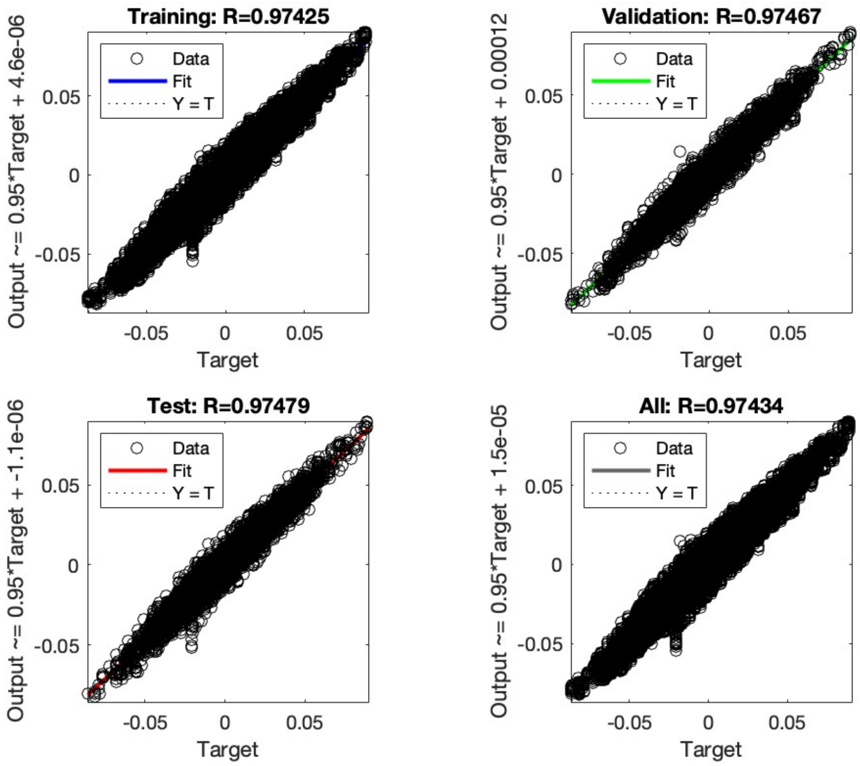
Appendix 20 MatLab Result 75-15-10-10



Appendix 21 Analysis Result 80-10-10-5

	Observations	MSE	R
Training	58982	3.2900e-05	0.9734
Validation	7373	3.2217e-05	0.9743
Test	7373	3.1953e-05	0.9737

Appendix 22 MatLab Result 80-10-10-5



Appendix 23 Analysis Result 80-10-10-10

	Observations	MSE	R
Training	58982	3.1788e-05	0.9742
Validation	7373	3.1572e-05	0.9747
Test	7373	3.1448e-05	0.9748

Appendix 24 MatLab Result 80-10-10-10